(19) PATENT BUREAU OF JAPAN (JP)
(12) OFFICIAL GAZETTE FOR GRANTED PATENTS (A)

(11) Japanese Patent Application Publication Kokai: Sho 63-500636
(43) Publication Date: March 10, 1988
Request for Examination: Not requested
Request for Preliminary Examination: Not requested (Total of 27 pages)
Department (cl.) 1 (1)

(51) Int. Cl.[4]	JP Class.	Intra-bureau Registration No.
C 12 P 21/02		C-6712-4B
A 61 K 37/02	ABE	
	ACC	
	ADA	
	ADT	

(54) Title of Invention: Production of pluripotent granulocyte colony-stimulating factor

(21) Patent Application No.: Patent Application No. Sho 61-504671

(86)(22) Application Date: August 22, 1986

(85) Translation Submission Date: April 23, 1987

(86) International Application: PCT/US86/01708

(87) International Publication No.: WO87/01132

(87) International Publication Date: February 26, 1987

Priority Claims: (32) August 23, 1985 (33) USA (US) (31) 768959
(32) March 3, 1986 (33) USA (US) (31) 835548

(72) Inventors: SOUZA, Lawrence M
1302 Witherspoon, Thousand Oaks, CA 91360 (US)

(71) Applicant: Kirin-Amgen, Inc.
1900 Oak Terrace Lane, Thousand Oaks, CA 91360 (US)

(74) Agent: Patent Attorney Kazuo Sato, and 2 others

(81) Designated States: AT (European Patent), AU, BE (European Patent), CH (European Patent), DE (European Patent), DK, FI, FR (European Patent), GB (European Patent), IT (European Patent), JP, KR, LU (European Patent), NL (European Patent), NO, SE (European Patent)

To be continued on the last page

Patent Claims

1. A purified and isolated polypeptide, characterized by the fact that it has a part or all of the primary structure and one or more of the biological properties of a naturally occurring pluripotent granulocyte colony-stimulating factor and is the product of prokaryotic or eukaryotic cell expression of an exogenous DNA sequence.

2. The polypeptide described under Patent Claim 1, characterized by the fact that it is free of association with any mammalian protein.

3. The polypeptide described under Patent Claim 1, characterized by the fact that the exogenous DNA sequence is a cDNA sequence.

4. The polypeptide described under Patent Claim 1, characterized by the fact that the exogenous DNA sequence is an artificial DNA sequence.

5. The polypeptide described under Patent Claim 1, characterized by the fact that the exogenous DNA sequence is a genomic DNA sequence.

6. The polypeptide described under Patent Claim 1, characterized by the fact that the exogenous DNA sequence is carried on an autonomously replicating DNA plasmid or viral vector.

7. The polypeptide described under Patent Claim 1, which contains a part or all of the primary structure of the human pluripotent granulocyte colony-stimulating factor shown in Table VII or any naturally occurring allelic variant thereof.

8. The polypeptide described under Patent Claim 1, which has the immunological properties of the naturally occurring pluripotent granulocyte colony-stimulating factor shown in Table VII or any naturally occurring allelic variant thereof.

9. The polypeptide described under Patent Claim 1, which has the in vitro biological activity of a naturally occurring pluripotent granulocyte colony-stimulating factor shown in Table VII or any naturally occurring allelic variant thereof.

10. The polypeptide described under Patent Claim 1, which is covalently bonded to a detectable label substance.

11. The polypeptide described under Patent Claim 1, wherein the said detectable label is a radiolabel.

12. A DNA sequence characterized by the fact that the DNA sequence is for use in securing the expression in a prokaryotic or eukaryotic host cell of a polypeptide product having at least a part of the primary structure and one or more of the biological properties of a naturally occurring pluripotent granulocyte colony-stimulating factor, and is selected from the following:

(a) the DNA sequence shown in Tables VI and VII or their complementary strands;

(b) DNA sequences which hybridize with the DNA sequence defined in (a) or fragments thereof; and

(c) DNA sequences which would hybridize with the DNA sequences defined in (a) and (b) if not for the degeneracy of the genetic codes.

13. A prokaryotic or eukaryotic host cell transformed or transfected with the DNA sequence described under Patent Claim 12 in a manner allowing the host cell to express the polypeptide product.

14. A polypeptide product of the expression of the DNA sequence described under Patent Claim 12 in a prokaryotic or eukaryotic host.

15. A purified and isolated DNA sequence for expression in a prokaryotic or eukaryotic host of a polypeptide product having a part or all of the primary structure and one or more of the biological properties of a naturally occurring pluripotent granulocyte colony-stimulating factor.

16. The DNA sequence described under Patent Claim 15.

17. The genomic DNA sequence described under Patent Claim 16.

18. The artificial DNA sequence described under Patent Claim 15.

19. The artificial DNA sequence described under Patent Claim 18, which includes one or more codons preferred for expression in E. coli.

20. The artificial DNA sequence described under Patent Claim 18, which codes for expression of the human pluripotent granulocyte colony-stimulating factor.

21. The artificial DNA sequence described under Patent Claim 20, which includes one or more codons preferred for expression in yeast cells.

22. The DNA sequence described under Patent Claim 16, 17 or 18, which codes for human pluripotent granulocyte colony-stimulating factor.

23. The DNA sequence described under Patent Claim 15, which is covalently associated with a detectable label substance.

24. The DNA sequence described under Patent Claim 23, wherein the detectable label is a radiolabel.

25. The single-stranded DNA sequence described under Patent Claim 23.

26. A DNA sequence coding for a polypeptide fragment or polypeptide analog of a naturally occurring pluripotent granulocyte colony-stimulating factor.

27. A DNA sequence coding for [Ala^{1}] hpG-CSF.

28. A biologically functional plasmid or viral DNA vector containing the DNA sequence described under Patent Claim 12, 15 or 26.

29. A prokaryotic or eukaryotic host cell stably transformed or transfected with the DNA vector described under Patent Claim 28.

30. A polypeptide product of the expression in a prokaryotic or eukaryotic host cell of the DNA sequence described under Patent Claim 15 or 26.

31. A synthetic polypeptide having a part or all of the amino acid sequence shown in Table VII and one or more of the in vitro biological activities of a naturally occurring pluripotent granulocyte colony-stimulating factor.

32. A synthetic polypeptide having a part or all of the amino acid sequence shown in Table VII, a part or all of the secondary structure and the biological activities of human pluripotent granulocyte colony-stimulating factor.

33. A process for the production of a polypeptide having a part or all of the primary structure and one or more of the biological properties of a naturally occurring pluripotent granulocyte colony-stimulating factor, comprising growing, under appropriate nutrient conditions, prokaryotic or eukaryotic host cells transformed or transfected with the DNA vector described under Patent Claim 28, and isolating the target polypeptide product expressed by the DNA sequence in the said vector.

34. Purified and isolated human pluripotent granulocyte colony-stimulating factor in a glycosylated or non-glycosylated form, free of association with any other human protein.

35. A pharmaceutical composition comprising an effective amount of the polypeptide described under Patent Claim 1 or 34 and a pharmaceutically acceptable diluent, adjuvant or carrier.

36. A method for providing a hematopoietic therapy to a mammal, comprising administering an effective amount of the polypeptide described under Patent Claim 1 or 34.

37. A method for arresting proliferation of leukemic cells, comprising administering an effective amount of the polypeptide described under Patent Claim 1 or 34.

38. A DNA sequence coding for [Ser^{17}] hpG-CSF.

39. A polypeptide product of the expression in a prokaryotic or eukaryotic host cell of the DNA sequence described under Patent Claim 38.

40. A biologically functional plasmid or viral DNA vector containing the DNA sequence described under Patent Claim 38.

41. A prokaryotic or eukaryotic host cell stably transformed or transfected with the DNA vector described under Patent Claim 40.

42. A DNA sequence coding for an analog of hpG-CSF selected from a group consisting of the following:

(Ala^{1}) hpG-CSF;

(Ser^{36}) hpG-CSF;

(Ser^{42}) hpG-CSF;

(Ser^{64}) hpG-CSF;

(Ser^{74}) hpG-CSF;

(Met^{-1}, Ser^{17}) hpG-CSF;

(Met^{-1}, Ser^{36}) hpG-CSF;

(Met^{-1}, Ser^{42}) hpG-CSF;

(Met^{-1}, Ser^{64}) hpG-CSF; and

(Met^{-1}, Ser^{74}) hpG-CSF;

43. A polypeptide product of the expression in a prokaryotic or eukaryotic host cell of the DNA sequence described under Patent Claim 42.

44. A preparation of hpG-CSF which is greater than 95% pure and contains less than 0.5 ng of pyrogen per 0.5 ng of hpG-CSF.

Detailed Description

Production of pluripotent granulocyte colony-stimulating factor

This is a continuation in-part of the co-pending US Patent Application No. 768,959 filed on August 23, 1985.

Technical Background

The present invention is related to hematopoietic growth factors and to polynucleotides coding for such factors in general. Specifically, the present invention is related to mammalian pluripotent granulocyte colony-stimulating factors, especially to the human pluripotent granulocyte colony-stimulating factor (hpG-CSF), their fragments and polypeptide analogs, and polynucleotides coding for all of them.

The human blood formation (hematopoietic) system supplies various white blood cells (including neutrophils, macrophages, basophils and mast cells), red blood cells (erythrocytes) and clot-forming cells (megakaryocytes and platelets). It has been estimated that the

hematopoietic system of an average male produces granulocytes and erythrocytes of the order of 4.5×10^{11} per year, equivalent to the total body weight [Dexter et al., BioEssays 2, 154-158 (1985)].

It is considered that specific hematopoietic growth factors in small amounts are the reasons for the differentiation of a small number of progenitor "stem cells" into the various blood cell lines, the dramatic proliferation of these lines, and the ultimate differentiation into mature blood cells from these lines. Because the hematopoietic growth factors are present in extremely small amounts, their detection and identification have relied on analytical techniques, which currently can only distinguish among the different factors based on their stimulating effects on cells cultured under artificial conditions. Consequently, there have been a large number of names for a much smaller number of such factors, which has been confusing. For example, IL-3, BPA, multi-CSR, HCGF, MCGF and PSF are all the abbreviations of the names for one single mouse hematopoietic growth factor [Metcalf, Science, 229, 16-22 (1985)]. Moreover, various mouse growth regulating glycoproteins have been reviewed [Burgess et al., J. Biol. Chem. 252, 1988 (1977); Das et al., Blood 58, 600 (1980); Ihle et al., J. Immunol. 129, 2431 (1982); Nicola et al., J. Biol. Chem. 258, 9017 (1983); Metcalf et al., Int. J. Cancer, 30, 773 (1982); and Burgess et al., Int. J. Cancer 26, 647 (1980)].

The application of recombinant genetic technology has established some order in this chaotic situation. For example, the amino acid and DNA sequences of human erythropoietin, which stimulates the production of erythrocytes, have been determined (Lin, PCT Application Publication No. 85/02610, published on June 20, 1985). Recombinant techniques have also been applied to the isolation of a cDNA for human granulocyte-macrophage colony-stimulating factor [Lee et al., Proc. Natl. Acad. Sci. USA 82, 4360-4364 (1985); Wong et al., Science 228, 810-814 (1985)]. Cloning of mouse genes has been described [Yokota et al., Proc. Natl. Acad. Sci. USA 81, 1070 (1984); Fung et al., 307, 233 (1984); Gough, et al., Nature 309, 763 (1980)]. In addition, human M-CSF has been described [Kawasaki et al., Science 230, 291 (1985)].

A human hematopoietic growth factor called human pluripotent granulocyte colony-stimulating factor (hpCSF) or pluripoietin has been shown to be present in the culture medium of a human bladder cancer cell line. This cell line is called 5637 and has been deposited with the American Type Culture Collection (Rockville, Maryland) under restrictive conditions with a deposit No. HTB-9. The hpCSF purified from this cell line has been reported to stimulate the proliferation and differentiation of pluripotent progenitor cells, leading to the production of all major blood cell types in assays using human bone marrow progenitor cells [Welte et al., Proc.

Natl. Acad. Sci. USA 82, 1526-1530 (1985)]. For the purification of the hpCSF, the following were used: $(NH_4)_2SO_4$ precipitation, anion exchange chromatography (DEAE-Cellulose, DE52), gel filtration (AcA54 column), and C18 reverse phase high performance liquid chromatography. A protein identified as hpCSF, which was eluted as the second activity peak on the C18 reverse phase HPLC, was reported to have a molecular weight (MW) of 18,000 as determined by sodium dodecyl sulfate (SDS)-polyacrylamide gel electrophoresis (PAGE) using silver staining. Early on hpCSF was reported to have an isoelectric point of 5.5 [Welte et al., J. Cell. Biochem. Suppl. 9A, 116 (1985)] and a potent differentiation activity for the mouse myelomonocytic leukemic cell line WEHI-3B D^+ [Welte et al., UCLA Symposia on Molecular and Cellular Biology, Gale et al. eds., New Series, 28 (1985)]. Preliminary studies have shown that the factor identified as hpCSF has a significant granulocyte colony-stimulating activity during the first seven days in a human CFU-GM assay.

Another factor called human CSF-β has also been isolated from the human bladder cancer cell line 5637 and shown to be a competitor in the binding of ^{125}I-labeled mouse granulocyte colony-stimulating factor (G-CSF) to WEHI-3B D^+ cells, in a dose-response relationship identical to that of unlabeled mouse G-CSF [Nicola et al., Nature 314, 625-628 (1985)]. This dose-response relationship had been specific before for the unlabeled mouse G-CSF, but not for factors such as M-CSF, GM-CSF or multi-CSF [Nicola et al., Proc. Natl. Acad. Sci. USA 81, 3765-3769 (1984)]. CSF-β and G-CSF are also unique among CSFs in that they both have a potent differentiation activity for WEHI-3B D^+ cells [Nicola et al., Immunology Today 5, 76-80 (1984)]. G-CSF stimulates mixed granulocyte/macrophage colony-forming cells at high concentrations [Nicola et al., ibid. (1984)], which is consistent with preliminary results showing the appearance of granulocytic, monocytic mixed granulocytic/monocytic and eosinophilic colonies (CFU-GEMM) after 14-day incubation of human bone marrow culture with hpCSF. CSF-β also has been shown to stimulate the formation of neutrophilic granulocytic colonies in an assay using mouse bone marrow cells. This property has been a criterion for identification of factors as a G-CSF. Accordingly, based on these similarities, human CSF-β has been identified as a G-CSF (granulocyte colony-stimulating factor) [Nicola et al., Nature 314, 625-628 (1985)].

Based on their common properties, it appears that the human CSF-β of Nicola et al. and the hpCSF of Welte et al. are the same factor that can be appropriately called human pluripotent granulocyte colony-stimulating factor (hpG-CSF). Characterization and recombinant production of hpG-CSF would be particularly desirable based on the reported activity of mouse G-CSF of completely inhibiting established transplanted myeloid leukemia cell line WEHI-$3B^+$ in mice

[Metcalf, Science, 229, 16-22 (1985)]. See also Sachs, Scientific American, 2848 (1), 40-47 (1986).

hpG-CSF may prove to be therapeutically significant and hence needs to be available in commercial scale quantities. However, its isolation from cell cultures is unlikely to provide adequate quantities of the material. Moreover, it is noteworthy, for example, that there appear restrictions on commercial use of cells from human tumor banks, such as the human bladder cancer cell line 5637 (A. T. C. C. HTB9) that has been reported as a source of isolation of natural hpCSF by Welte et al. (1985, ibid.).

Summary of the Invention

By the present invention, DNA sequences coding for all or a part of hpG-CSF are provided. Such sequences may include: the incorporation of codons "preferred" for expression by selected non-mammalian hosts; the provision of sites for cleavage by restriction endonuclease enzymes; and the provision of additional initial, terminal or medial DNA sequences which facilitate construction of readily expressible vectors. The present invention also provides DNA sequences coding for microbial expression of polypeptide analogs or derivatives of hpG-CSF which differ from naturally occurring forms in terms of the identity or location of one or more amino acid residues (that is, deletion analogs containing less than all of the specific residues of hpG-CSF; substitution analogs, such as [Ser^{17}]hpG-CSF, wherein one or more specific residues are substituted by other residues, and addition analogs wherein one or more amino acid residues is added to a terminal or medial portion of the polypeptide), and which share some or all the properties of naturally occurring forms.

Novel DNA sequences of the present invention include sequences useful in securing expression in prokaryotic or eukaryotic host cells of polypeptide products having at least a part of the primary structure and one or more of the biological properties of the naturally occurring pluripotent granulocyte colony-stimulating factor. DNA sequences of the present invention specifically include: (a) the DNA sequence shown in Table VII and Table VIII or their complementary strands; (b) a DNA sequence which can hybridize (under hybridization conditions described herein or more stringent conditions) to the DNA sequences in Table VII or to their fragments; and (c) DNA sequences which can hybridize to the DNA sequences in Table VII if not for the degeneracy of the genetic codes. In particular, (b) includes genomic DNA sequences coding for allelic variants of hpG-CSF and /or the pluripotent granulocyte colony-stimulating factors of other mammalian species. Particularly included in (c) are artificial DNA sequences coding for hpG-CSF, fragment of hpG-CSF and analogs of hpG-CSF. These DNA

sequences may incorporate codons facilitating translation of messenger RNA in microbial hosts. Such artificial sequences may readily be constructed according to the method of Alton, et al. (PCT Published Application No. WO83/04053).

The present invention also covers polypeptides coded for by portions of a DNA complementary to the top strand of the human cDNA or genomic DNA sequence shown in Tables VII or VIII, that is, "complementary inverted proteins" as described by Tramontano et al. [Nucleic Acids Res., 12, 5049-5059 (1984)].

The present invention provides purified and isolated polypeptide products having part or all of the primary structure (that is, a continuous sequence of amino acid residues) and one or more of the biological properties (e.g, immunological properties and in vitro biological activities) and physical properties (e.g. molecular weight) of the naturally occurring hpG-CSF including its allelic variants. These polypeptides are also characterized by the fact that they are the products of chemical synthesis procedures or of prokaryotic or eukaryotic host expression (e. g. by bacterial, yeast, higher plant, insect and mammalian cells in culture) of exogenous DNA sequences obtained by genomic or cDNA cloning or by gene synthesis. The products of typical yeast (e. g. Saccharomyces cerevisiae) or prokaryote [e. g. Escherichia coli (E. coli)] host cells are free of association with any mammalian proteins. The products of expression in vertebrate (e. g. non-human mammalian and avian) cells are free of association with any human proteins. Depending on the host used, polypeptides of the present invention may be glycosylated with mammalian or other eukaryotic carbohydrate chains or may be non-glycosylated. Polypeptides of the present invention may also include an initial methionine residue (at position -1).

The present invention also covers a pharmaceutical composition useful for hpG-CSF therapy, comprising an effective amount of the polypeptide product of the present invention along with an appropriate diluent, adjuvant and/or carrier.

The polypeptide product of the present invention can be "labeled" with a detectable marker substance (e. g. radiolabeled with ^{125}I) to provide a reagent for detection and quantification of hpG-CSF in tissue and fluid samples such as blood and urine. The DNA products of the present invention can also be labeled with a detectable marker (e. g. radiolabels and non-radioisotope labels such as biotin) for DNA hybridization in order to locate hpG-CSF gene position and/or the position of any other related gene family member in a chromosomal map. They can also be used to confirm hpG-CSF gene disorders at the level of DNA and as a gene marker for confirming neighboring genes and their disorders.

The polypeptide product of the present invention can be useful, alone or in combination with other hematopoietic growth factors or drugs for the treatment of hematopoietic disorders, such as aplastic anemia. They can also be useful for the treatment of hematopoietic disorders caused by chemotherapy or radiation therapy. For example, the success of bone marrow transplantation can be further increased by use of hpG-CSF. Use of hpG-CSF may also be useful for the treatment of burn and bacterial inflammation and for wound healing. In addition, hpG-CSF may also be useful for the treatment of leukemia, based on its reported activity for leukemia cell differentiation [Welte et al., Proc. Natl. Acad. Sci. USA 82, 1526-1530 (1985); and Sachs, ibid.].

Many advantages and aspects of the present invention will be apparent to those skilled in this field after studying the following detailed description, which describes the currently preferred embodiments of the present invention using practical examples.

Brief Description of the Drawings

The drawing shows a partial restriction endonuclease map of the hpG-CSF gene, with arrows indicating the sequencing strategy used to obtain the genomic sequences.

Detailed Description

By the present invention, DNA sequences coding for a part or all of the polypeptide sequence of hpG-CSF were isolated and characterized.

The following examples are provided to describe the present invention in detail, in particular on procedures used prior to the identification of hpG-CSF cDNA and genomic DNA, procedures leading to the identification and their sequencing, construction of the expression systems using the cDNA, genomic DNA and artificial genes, and verification of expressed hpG-CSF and its analog products in these systems.

More specifically, Practical Example 1 is related to amino acid sequence determination of hpG-CSF. Practical Example 2 is related to preparation of a cDNA library for screening by colony hybridization. Practical Example 3 is related to construction of probes for hybridization. Practical Example 4 is related to screening by hybridization, identification of positive clones, DNA sequencing of a positive cDNA clone and generation of information on primary structure (amino acid sequence) of the polypeptide. Practical Example 5 is related to identification and sequencing of a genomic clone coding for hpG-CSF. Practical Example 6 is related to

construction of an artificial gene coding for hpG-CSF where codons preferred for E. coli expression were used.

Practical Example 7 is related to procedures for constructing an E. coli transformation vector containing a hpG-CSF-encoding DNA, use of the vector in prokaryotic expression of hpG-CSF and analysis of properties of the recombinant product of the present invention. Practical Example 8 is related to procedures for generating hpG-CSF analogs whose cysteine residues were substituted by other appropriate amino acid residues by mutagenesis of the hpG-CSF-encoding DNA. Practical Example 9 is related to procedures for constructing a vector containing an hpG-CSF analog-encoding DNA derived from the positive clone, use of the vector to transfect COS-1 cells, and growth of the transfected cells by culture. Practical Example 10 is related to physical and biological properties of the recombinant polypeptide product of the present invention.

Practical Example 1

(A) Sequencing of material obtained by literature method

A sample (3-4 μg, 85-90% pure on SDS-PAGE by silver staining) of hpG-CSF was obtained from the Sloan Kettering Institute (New York, New York), isolated and purified as described by Welte et al. [Proc. Natl. Acad. Sci. USA 82, 1526-1530 (1985)].

The N-terminal amino acid sequence of this hpG-CSF sample was determined in Run #1 by microsequence analysis using an AB407A gas phase sequencer (Applied Biosystems, Foster City, California) to obtain the sequence data shown in Table 1. In Tables I-IV single letter codes are used, with X denoting a residue which was not unambiguously determined. Residues in parentheses were only determined tentatively or alternatively.

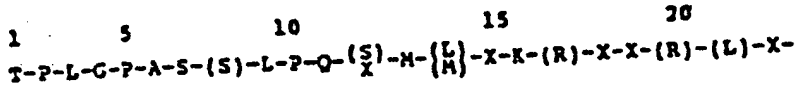
Table I

1				5				10				15			
K	P	L	G	P	A	S	K	L	K	Q	(G,V,S)	G	L	X	X	X

There was a high level of background in every cycle of the run, for which the results are shown in Table I, indicating that the sample contained many contaminating components, probably being chemical residues from purification. The sequence was retained only for reference purpose.

In Run #2, a second sample (5-6 μg, 95% pure) obtained from Sloan Kettering as for Run #1 was subjected to the same sequencing procedures as in Run #1. This sample was from the same lot of material used to generate Fig. 4 of the Welte et al.'s paper [Proc. Natl. Acad. Sci. USA 82, 1526-1530 (1985)].

Table II



Although more residues were determined, Run #2 still did not provide a long enough and unambiguous sequence, from which an appropriate number of probes could be prepared to search for hpG-CSF DNA. It was estimated that at least 1536 probes would have been needed to attempt isolation of a cDNA based on the sequence data shown in Table II. Again, contamination of the sample was considered to be the problem.

Accordingly, a third sample (3-5 μg, 40% pure) was obtained from Sloan Kettering as above. This sample was electroblotted after SDS-PAGE separation for further purification. However, sequence analysis of this sample yielded no useful data.

(B) Sequencing of material obtained by the revised method

To obtain a sufficient amount of pure material to perform definitive amino acid sequence analysis, cells of the bladder cancer cell line 5637 (subclone 1A6) produced as at Sloan Kettering were obtained from Dr. E. Platzer. The cells were initially cultured in flasks to confluence in Iscove's medium (GIBCO, Grand Island, New York). After reaching confluence, the cells were trypsinized and seeded into roller bottles (1.5 flask/bottle) each containing 25 mL of pre-conditioned Iscove's medium with 5% CO_2. The cells were cultured overnight at 37°C at 0.3 rpm.

Cytodex-1 beads (Pharmacia, Uppsala, Sweden) were washed and sterilized by the following procedures. 8 g of the beads were placed in a bottle, and 400 mL of PBS was added. The beads were suspended by gentle agitation for 3 hr. After allowing the beads to settle, the PBS was removed, and the beads were rinsed with PBS. Fresh PBS was added, and then the beads were autoclaved for 15 min. Before use, the beads were washed in Iscove's medium containing 10% fetal calf serum (FCS), followed by addition of fresh medium containing 10% FCS.

All but 30 mL of the medium was removed from each bottle, and 30 mL of fresh medium containing 10% FCS and 40 mL of the bead suspension were added to each bottle. The bottles were treated with 5% CO_2 and bubbles were removed by suction. The bottles were placed on roller racks at 3 rpm for 0.5 hr, and then the speed was reduced to 0.3 rpm. After 3 hr, an additional flask of cells was trypsinized and added to each roller bottle containing beads.

At a confluence of 40-50%, the cells under the roller bottle culture were washed with 50 mL of PBS and rolled for 10 min. The PBS was removed, and then the cells were cultured for 48 hr in medium A (Iscove's medium containing 0.2% FCS, 10^{-8} M hydrocortisone, 2 mM glutamine, 100 unit/mL penicillin, and 100 µg/mL streptomycin). Subsequently, the culture was subjected to centrifugation at 3000 rpm for 15 min, and the supernatant was stored at –70°C. The cells were cultured again in medium A containing 10% FCS for 48 hr. The medium was discarded, and the cells were washed with PBS as above, followed by culture in medium A for 48 hr. The supernatant was collected again and treated as above.

About 30 L of the medium from 1A6 cell culture was concentrated to about 2 L using a Millipore Pellicon unit equipped with 2 cassettes having a cutoff of 10,000 MW at a filtration rate of about 200 mL/min and at a retentate rate of about 1000 mL/min. The concentrate was diafiltered with about 10 L of 50 mM Tris buffer (pH 7.8) using the same apparatus at the same flow rates. The diafiltered concentrate was applied at 40 mL/min to a 1 L DE Cellulose column equilibrated in 50 mM Tris buffer (pH 7.8). The column was washed at the same flow rate with 1 L of 50 mM Tris buffer (pH 7.8) and then with 2 L of 50 mM Tris buffer (pH 7.8) containing 50 mM NaCl. The column was then sequentially eluted with six 1-L solutions of 50 mM Tris buffer (pH 7.5) containing NaCl at the following concentrations: 75 mM, 100 mM, 125 mM, 150 mM, 200 mM and 300 mM. Then 50-mL fractions were collected. The active fractions were pooled and concentrated to 65 mL on an Amicon ultrafiltration unit with stirrer, equipped with a YM5 membrane. The concentrate was applied to a 2 l AcA54 gel filtration column equipped equilibrated in PBS. The column was eluted at a flow rate of 80 mL/hr, and 10-mL fractions were collected. The active fractions were pooled and applied directly onto a C4 column in a high performance liquid chromatography (HPLC) system.

Samples in volumes of 125-850 mL containing 1-8 mg protein, about 10% of which was hpG-CSF, were applied to the column at a flow rate of 1-4 mL/min. The column was washed with 0.1 M ammonium acetate (pH 6.0-7.0) in 80% 2-propanol at a flow rate of 1 mL/min. Then 1-mL fractions were collected, and the proteins were monitored by measuring absorbance at 220 nm, 260 nm and 280 nm.

By the purification, fractions containing hpG-CSF were clearly separated (as fraction Nos. 72 and 73 among 80 fractions) from other protein-containing fractions. hpG-CSF was isolated (150-300 μg) with a purity of about 80±5% and with a yield of about 50%. From this purified material, 9 μg was used in Run #4 for an amino acid sequence analysis. The protein sample was applied to a TFA-activated glass fiber disc without polybrene. Sequence analysis was carried out with an AB470A sequencer according to the methods of Hewick et al. [J. Biol. Chem., 256, 7990-7997 (1981)] and Lai [Anal. Chim. Acta, 163, 243-248 (1984)]. The result of Run #4 is shown in Table III.

Table III

```
1                                   5                                  10
Thr - Pro - Leu - Gly - Pro - Ala - Ser - Ser - Leu - Pro-

                        15                                  20
Gln - Ser - Phe - Leu - Leu - Lys -(Lys)- Leu -(Glu)- Glu-

                                    25                                  30
Val - Arg - Lys - Ile -(Gln)- Gly - Val - Gly - Ala - Ala-

Leu - X   - X  -
```

In Run #4, beyond 31 cycles (corresponding to residue 31 in Table III), no further significant sequence information was obtained. In order to obtain a longer unambiguous sequence, in Run #5, 14 μg of hpG-CSF purified from the conditioned medium was reduced with 10 mL of 2-mercaptoethanol for 1 hr at 45°C, and then thoroughly dried in a vacuum. The protein residue was then redissolved in 5% formic acid and applied to a polybrenized glass fiber disc. Sequence analysis was carried out as in above Run #4. The results of Run #5 are shown in Table IV.

Table IV

```
1                                   5                                  10
Thr - Pro - Leu - Gly - Pro - Ala - Ser - Ser - Leu - Pro -
                                    15                                  20
Gln - Ser - Phe - Leu - Leu - Lys - Cys - Leu - Glu - Gln-

                                    25                                  30
Val - Arg - Lys - Ile - Gln - Gly - Asp - Gly - Ala - Ala -
                                    35                                  40
Leu - Gln - Phe - Lys - Leu - Gly - Ala - Thr - Tyr - Lys -
                                45
Val - Phe - Ser - Thr -(Arg)-(Phe)-(Met)-X-
```

The amino acid sequence shown in Table IV was sufficiently long (44 residues) and unambiguous to prepare probes for obtaining hpG-CSF cDNA as described below.

Practical Example 2

Among standard procedures for isolating cDNA sequences of interest, there is a method in which a plasmid-borne cDNA "library" is derived from reverse transcription of mRNA abundant in appropriate donor cells, selected on the basis of their expression of a target gene. If substantial portions of the amino acid sequence of a polypeptide are known, a labeled, single-stranded DNA probe sequence duplicating a sequence putatively present in the "target" cDNA may be used in DNA/DNA hybridization procedures carried out on cloned copies of the cDNA denatured to the single stranded form [Weissman et al., U.S. Patent No. 4,394,443; Wallace et al., Nucleic Acids Res., 6, 3543-3557 (1979); Reyes et al., Proc. Natl. Acad. Sci. (USA), 79, 3270-3274 (1982); and Jaye et al., Nucleic Acids Res., 11, 2325-2335 (1983)]. Also, see U.S. Patent No. 4,358,535 (Falkow et al.) for DNA/DNA hybridization procedures in diagnostic procedures and Davis et al.'s "A Manual for Genetic Engineering, Advanced Bacterial Genetics", Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1980) at pp. 55-58 and 174-176 for colony and plaque hybridization techniques.

The total RNA was extracted from approximately 1 g of cells from the bladder cancer cell line 5637 (IA6) using a guanidinium thiocyanate procedure for quantitative isolation of intact RNA [Chirgwin et al., Biochemistry, 18, 5294-5299 (1979)].

The sterile aqueous RNA solution contained total RNA from the IA6 cells. To obtain only the messenger RNA from the total RNA solution, the solution was passed through a column containing oligodeoxythymidylate [oligo(dT)] (Collaborative Research, Inc., Waltham, Massachusetts). Poly-Adenylated (poly-A^+) tails characteristic of messenger RNA adhered to the column while ribosomal RNA was eluted out. As a result of this procedure, approximately 90 μg of poly-adenylated messenger RNA (poly-A^+ mRNA) was isolated. The isolated poly-A^+ messenger RNA was pre-treated with methylmercury hydroxide (Alpha Ventron, Danvers, Massachusetts) at a final concentration of 4 mM for 5 min at room temperature prior to use in a cDNA reaction. The methylmercury hydroxide treatment suppresses interactions between messenger RNAs and with contaminating molecules, which inhibit translation [Payvar et al., J Biol. Chem., 258, 7636-7642 (1979)].

According to the Okayama procedure [Okayama et al., Molecular and Cellular Biology, 2, 161-170 (1982)], a cDNA bank was prepared using the mRNA obtained IA6 cells. The

cDNAs were then transformed, by incubation, into the host microorganism E. coli K-12 strain HB101 for amplification.

Practical Example 3

Hybridization probes designed on the basis of the amino terminal sequence of hpG-CSF, shown in Table VI, consisted of a set of 24 oligonucleotides each being 23 bases in length and containing three inosine residues. The oligonucleotide probes were synthesized according to the procedure of Caruthers et al. [Genetic Engineering, 4, 1-18 (1982)] and labeled with γ-^{32}P ATP by kinasing with polynucleotide kinase. The oligonucleotide probes corresponding to messenger RNA for residues 23-30 of the sequence of Table IV are shown in Table V.

Table V

hpG-CSF probe

5' GC IGC ICC A/C TC ICC T/C TC G/A/T AT T/C TT 3'

The assignment of neutrality to I's was based on the published work of Takahashi et al. [Proc. Natl. Acad. Sci. (USA), 82, 1931-1935 (1985)]. However, inosine may have a destabilizing effect if it is base-paired with a G or T [Ohtsuka, et. al., J. Biol. Chem, 260, 2605-2608 (1985). In Takahashi et al.'s study, inosines might appear to have a neutral effect because they averaged out as a group to near neutrality (e.g., three being paired favorably with C and two not favorably with T).

To test the effect of having I's base-paired with G's, control experiments were designed using N-myc gene sequences and a clone. The sequences selected from the N-myc gene had the same overall G and C content at the first two positions of each codon as was defined by the hpG-CSF probes. Thus, the test probes of N-myc were of the same length, contained I's in the same relative positions and had potentially the same average Tm (62-66°C, not due to the 3 or 4 inosine residues included) as the hpG-CSF probes.

Two sets of test probes of N-myc were prepared according to the procedure of Caruthers et al. (ibid.). Set I, as shown in Table VI, included the following: 1, a 23-mer with perfect match; 2, 3 third position C's were replaced with I's representing the worst possible case of adding I's; and 3, 4 third position C's were replaced with I's. Set II of test probes was designed to represent a more random distribution of inosine base pairs, which might give an overall neutral base pairing effect. Set II, as shown in Table VI, included the following:

4, containing two I's that were base-paired with C's and one with a G; and 5, identical to 4 with the addition of one more I:G base pair.

Table VI

N-myc test probes

1. 5'CAC AAC TAT GCC GCC CCC TCC CC3'

2. 5'CAC AAC TAT GCI GCC CCI TCI CC3'

3. 5'CAI AAC TAT GCI GCC CCI TCI CC3'

4. 5'AAC GAG CTG TGI GGC AGI CCI GC3'

5. 5'AAI GAG CTG TGI GGC AGI CCI GC3'

Five replica filters containing N-myc DNA sequences and chicken growth hormone DNA sequences (as a negative control) were baked in a vacuum oven for 2 hr at 80°C prior to hybridization. All the filters were hybridized as described in Example 4 with the hpG-CSF probes except that the period of hybridization was only 6 hr. The filters were washed three times at room temperature then once at 45°C, 10 min each. The filters were monitored with a Geiger counter.

The filter with N-myc probe 3 gave a very weak signal relative to the other four probed filters and was not washed any further. After washing for 10 min at 50°C, the Geiger counter gave the following percent signals with probe 1 being normalized to 100%: probe 2, 20%; probe 3 (45°C), 2%; probe 4, 92%; and probe 5, 75%. After washing at 55°, the percentages were the following: probe 2, 16%; probe 4, 100%; and probe 5, 80%. A final washing at 60°C yielded the following percentages: probe 2, 1.6%; probe 4, 90%; and probe 5, 70%.

Thus in the presence of three I's, as with probes 2 and 4, up to a 60-fold difference in signal was observed, as the theoretical Tm (I's not included in the calculation) was approached [based upon a worst case of I base pairing (probe 2) and a relatively neutral case of I base pairing (probe 4)].

The standardization information obtained by the test hybridization of N-myc was utilized in washing and monitoring of the hpG-CSF hybridization described below, to gauge the degree of confidence with which results of lower stringent washing might be accepted.

Practical Example 4

According to the procedure of Hanahan et al. [J. Mol. Biol., 166, 557-580 (1983)], bacteria containing the recombinant with cDNA insert prepared in Practical Example 2 were spread on 24 nitrocellulose filters (Millipore, Bedford, Massachusetts) laid out on agar plates. The plates were then incubated, thereby obtaining approximately 150,000 colonies, which were then replica-plated to 24 other nitrocellulose filters. The replicas were incubated until distinct colonies appeared. The bacteria on the filters were lysed on a sheet of Whatman 3 MM paper slightly saturated with sodium hydroxide (0.5M) containing NaCl (1.5M) for 10 min. When the filters were nearly dry, they were baked for 2 hr at 80°C in a vacuum oven prior to nucleic acid hybridization [Wahl et al., Proc. Natl. Acad. Sci. (USA), 76, 3683-3687 (1979); and Maniatis et al., Cell, 81, 163-182 (1976)].

The filters were prehybridized for 2 hr at 65°C in 750 mL of 10x Denhardt's, 0.2% SDS and 6X SSC. The filters were rinsed in 6X SSC, and then four filters were placed in a bag, followed by hybridization for 14 hr in 6X SSC and 10X Denhardt's. There was approximately 15 mL of the solution per bag containing 50 x 10^6 cpm of ^{32}P-labeled probe (oligonucleotides).

After hybridization, the filters were washed three times in 6X SSC (1 L/wash) at room temperature for 10 min each. The filters were then washed at 45°C for 15 min using 1 L of 6X SSC. The filters were autoradiographed for 2 hr at -70°C using an intensifying screen and Kodak XAR-2 film. On this autoradiography, there were 40-50 positive signals detected including 5 very intense signals.

The areas containing the strongest five signals and an additional five positives were scraped from the master plates and replated for a secondary screening using the same probe mixture under the same conditions. The wash procedure differed in that the high temperature wash consisted of two washes at 55°C for 15 min each and then one at 60°C for 15 min. Based on the N-myc test probe study of Practical Example 3, the temperature for the final wash in the second screening was increased because the aggregate melting temperature for the 24 23-mers was 60-68°C, similar to that of the N-myc probes. After the second wash at 55°C, the filters were allowed to remain wet and were subjected to autoradiography. Comparison of this autoradiograph with a second autoradiograph taken for a similar period of time after a final wash at 60°C showed that only two of the 10 clones being tested did not have a substantial loss in signal after increasing the temperature from 55°C to 60°C. These two clones were later shown to have nearly identical lengths and restriction endonuclease patterns. One clone designated Ppo2, was selected for sequencing.

Sequencing of the recombinant hpG-CSF cDNA clone Ppo2 obtained by the above procedure was performed by the dideoxy method of Sanger et al. [Proc. Natl. Acad. Sci. (USA), 74, 5463-5467 (1977)]. The single-stranded DNA phage M-13 was used as a cloning vector for supplying single-stranded DNA templates from the double-stranded cDNA clones. The Sanger et al.'s method revealed the sequence shown in Table VII, along with its deduced amino acid sequence and a complementary strand in the polypeptide coding region.

Table VII

```
                       HindIII
      5' - AGCTTGGACTCAGCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGG[fIIIIIIII]

      -12     -10                                         -1  +1
          Leu Trp His Ser Ala Leu Trp Thr Val Gln Glu Ala Thr Pro Leu Gly Pro
          CTG TGG CAC AGT GCA CTC TGG ACA GTG CAG GAA GCC ACC CCC CTG GGC CCT
          GAC ACC GTG TCA CGT GAG ACC TGT CAC GTC CTT CGG TGG GGG GAC CCG GGA
                  HglAI                                           ApaI
                  10                                  20
Ala Ser Ser Leu Pro Gln Ser Phe Leu Leu Lys Cys Leu Glu Gln Val Arg Lys Ile Gln
GCC AGC TCC CTG CCC CAG AGC TTC CTG CTC AAG TGC TTA GAG CAA GTG AGG AAG ATC CAG
CGG TCG AGG GAC GGG GTC TCG AAG GAC GAG TTC ACG AAT CTC GTT CAC TCC TTC TAG GTC

                30                                  40
Gly Asp Gly Ala Ala Leu Gln Glu Lys Leu Cys Ala Thr Tyr Lys Leu Cys His Pro Glu
GGC GAT GGC GCA GCG CTC CAG GAG AAG CTG TGT GCC ACC TAC AAG CTG TGC CAC CCC GAG
CCG CTA CCG CGT CGC GAG GTC CTC TTC GAC ACA CGG TGG ATG TTC GAC ACG GTG GGG CTC

                50                                  60
Glu Leu Val Leu Leu Gly His Ser Leu Gly Ile Pro Trp Ala Pro Leu Ser Ser Cys Pro
GAG CTG GTG CTG CTC GGA CAC TCT CTG GGC ATC CCC TGG GCT CCC CTG AGC AGC TGC CCC
CTC GAC CAC GAC GAG CCT GTG AGA GAC CCG TAG GGG ACC CGA GGG GAC TCG TCG ACG GGG

                70                                  80
Ser Gln Ala Leu Gln Leu Ala Gly Cys Leu Ser Gln Leu His Ser Gly Leu Phe Leu Tyr
AGC CAG GCC CTG CAG CTG GCA GGC TGC TTG AGC CAA CTC CAT AGC GGC CTT TTC CTC TAC
TCG GTC CGG GAC GTC GAC CGT CCG ACG AAC TCG GTT GAG GTA TCG CCG GAA AAG GAG ATG

                90                                  100
Gln Gly Leu Leu Gln Ala Leu Glu Gly Ile Ser Pro Glu Leu Gly Pro Thr Leu Asp Thr
CAG GGG CTC CTG CAG GCC CTG GAA GGG ATC TCC CCC GAG TTG GGT CCC ACC TTG GAC ACA
GTC CCC GAG GAC GTC CGG GAC CTT CCC TAG AGG GGG CTC AAC CCA GGG TGG AAC CTG TGT
```

Table VII (Cont'd)

```
                  110                                          120
Leu Gln Leu Asp Val Ala Asp Phe Ala Thr Thr Ile Trp Gln Gln Met Glu Glu Leu Gly
CTG CAG CTG GAC GTC GCC GAC TTT GCC ACC ACC ATC TGG CAG CAG ATG GAA GAA CTG GGA
GAC GTC GAC CTG CAG CGG CTG AAA CGG TGG TGG TAG ACC GTC GTC TAC CTT CTT GAC CCT

                  130                                          140
Met Ala Pro Ala Leu Gln Pro Thr Gln Gly Ala Met Pro Ala Phe Ala Ser Ala Phe Gln
ATG GCC CCT GCC CTG CAG CCC ACC CAG GGT GCC ATG CCG GCC TTC GCC TCT GCT TTC CAG
TAC CGG GGA CGG GAC GTC GGG TGG CTC CCA CGG TAC GGC CGG AAG CGG AGA CGA AAG GTC

                  150                                          160
Arg Arg Ala Gly Gly Val Leu Val Ala Ser His Leu Gln Ser Phe Leu Glu Val Ser Tyr
CGC CGG GCA GGA GGG GTC CTG GTT GCC TCC CAT CTG CAG AGC TTC CTG GAG GTG TCG TAC
GCG GCC CGT CCT CCC CAG GAC CAA CGG AGG GTA GAC GTC TCG AAG GAC CTC CAC AGC ATG

                  170              174
Arg Val Leu Arg His Leu Ala Gln Pro OP
CGC GTT CTA CGC CAC CTT GCC CAG CCC TGA GCC AAG CCC TCC CCA TCC CAT GTA TTT ATC
CGC CAA GAT GCG GTG GAA CGG GTC GGG ACT

TCT ATT TAA TAT TTA TGT CTA TTT AAG CCT CAT ATT TAA AGA CAG GGA AGA GCA GAA CGG

AGC CCC AGG CCT CTG TGT CCT TCC CTG CAT TTC TGA GTT TCA TTC TCC TGC CTG TAG CAG
       StuI

TGA GAA AAA GCT CCT GTC CTC CCA TCC CCT GGA CTG GGA GGT AGA TAG GTA AAT ACC AAG

TAT TTA TTA CTA TGA CTG CTC CCC AGC CCT GGC TCT GCA ATG GGC ACT GGG ATG AGC CGC

TGT GAG CCC CTG GTC CTG AGG GTC CCC ACC TGG GAC CCT TGA GAG TAT CAG GTC TCC CAC
```

Table VII (Cont'd)

```
GTG GGA GAC AAG AAA TCC CTG TTT AAT ATT TAA ACA GCA GTG TTC CCC ATC TGG GTC CTT

GCA CCC CTC ACT CTG GCC TCA GCC GAC TGC ACA GCG GCC CCT GCA TCC CCT TGG CTG TGA

GGC CCC TGG ACA AGC AGA GGT GGC CAG AGC TGG GAG GCA TGG CCC TGG GGT CCC ACG AAT

TTG CTG GGG AAT CTC GTT TTT CTT CTT AAG ACT TTT GGG ACA TGG TTT GAC TCC CGA ACA

TCA CCG ACG TGT CTC CTG TTT TTC TGG GTG GCC TCG GGA CAC CTG CCC TGC CCC CAC GAG

GGT CAG GAC TGT GAC TCT TTT TAG GGC CAG GCA GGT GCC TGG ACA TTT GCC TTG CTG GAC

GGG GAC TCG GGA TGT GGG AGG GAG CAG ACA GGA GGA ATC ATG TCA GGC CTG TGT GTG AAA
                                                    StuI

GGA AGC TCC ACT GTC ACC CTC CAC CTC TTC ACC CCC CAC TCA CCA GTG TCC CCT CCA CTG

TCA CAT TGT AAC TGA ACT TCA GGA TAA TAA AGT GTT TGC CTC CA
```

[illegible line]

The following characteristics of the sequence of Table VII are significant. At the 5' end of the sequence there were bases corresponding to those of the poly G cDNA linker. This was followed by about five bases (designated as "N") whose sequence could not readily be determined ambiguously by Sanger et al.'s method due to the preceding multiple G's.

The sequence thereafter reveals a series of 12 codons encoding a portion of a putative leader sequence for the polypeptide. Based on the correspondence to the amino terminal sequence of the natural isolates of hpCSF described in Example 1, the initial threonine residue of the putative "mature" form of hpG-CSF is indicated by +1. Mature hpG-CSF is therefore revealed to include 174 residues as indicated. Following the "stop" codon (the OP codon, TGA) there are approximately 856 bases of an untranslated 3' sequence and multiple A's of the poly A "tail". Unique HgiAi, and ApaI restriction endonuclease recognition sites, as well as two StuI sites (discussed below with construction of procaryotic and eukaryotic expression systems) are also shown in Table VII. Due to the lack of asparagine residues in the polypeptide, there are no apparent sites for N-glycosylation. The underscored 6 bases near the end of the 3' untranslated sequence represent a potential polyadenylation site.

It is noteworthy that two additional cDNA clones identified by the hybridization procedures described above from among a total of 450,000 clones did not contain DNA encoding the entire leader sequence from the transcription initiation site onward. Indeed, all three hpG-CSF clones terminated in the 5' region at exactly the same site, indicating that the secondary structure of the mRNA transcribed severely hinders cDNA formation beyond this site. Thus, in practice, cDNA expression screening as described by Okayma et al. [Mol. and Cell. Biol. 3, 280-289 (1983)] and as actually used to isolate GM-CSF by Wong et al. [Science 228, 810-814 (1985)] could not have been readily applied to isolation of hpG-CSF DNA because such isolation normally relies on the presence of a full-length cDNA transcript in the clones assayed.

The above sequence is difficult to use directly for expression of hpG-CSF in a microbial host. To secure such expression, the hpG-CSF coding region should be provided with an initial ATG codon and the sequence should be inserted in a transformation vector at a site under control of an appropriate promoter/regulator DNA sequence.

Practical Example 5

In this practical example, the cDNA encoding hpG-CSF as isolated in the previous practical example was used to screen for a genomic clone. A phage lambda human fetal liver genomic library (prepared according to the procedures of Lawn et al., Cell, 15, 1157-1174 (1978) and obtained from T. Maniatis) was screened using a nick- translated probe consisting of two hpG-CSF cDNA fragments isolated by digestion with HgiAI and StuI (HgiAI to StuI, 649 bp; StuI to StuI, 639 bp). A total of approximately 500,000 phages were plated on 12 (15 cm) Petri dishes and, after plaques lifted, hybridized to the probe using the Benton/Davison procedures [Benton et al., Science, 196, 180 (1977)]. A total of 12 positive clones were identified. Three

clones (1-3) yielding the strongest signals on autoradiography in a secondary screening were grown in 1 culture and then mapped by restriction enzyme digestion and Southern blotting using a radiolabeled 24-mer oligonucleotide (kinased with γ-32P ATP) with the sequence 5' CTGCACTGTCCAGAGTGCACTGTG3'. Mapping results showed that clones 1 and 3 were identical and clone 2 contained 2000 additional bases on the 5'-end of the hpG-CSF gene. Therefore, clone 2 was used for further characterization. DNA from clone 2 was digested with R1 to release an 8500 bp hpG-CSF-containing fragment, which was subsequently subcloned into pBR322 and further mapped by restriction endonuclease digestion, Southern Blotting, M13 subcloning and sequencing. The sequences are shown in Table VIII.

Table VIII

```
GGGGACAGGCTTGAGAATCCCAAAGGAGAGGGGCAAAGGACACTGCCCCCCGCAAGTCTGCCAGAGCAGAGAGGGAGACCCCGACTCAGCTGCCACTTCCC   100

CACAGGCTCGTGCCGCTTCCAGGCGTCTATCAGCGGCTCAGCCTTTGTTCAGCTGTTCTGTTCAAACACTCTGGGGCCATTCAGGCCTGGGTGGGGCAGC   200

GCGAGGAAGGCAGTTTCAGGGGGGCAAGGCGACGTCAAAGGAGGATCAGAGATTCCACAATTTCACAAAACTTTCGCAAACAGCTTTTTGTTCCAACCCC   300

CCTGCATTGTCTTGGACACCAAATTTGCATAAATCCTGGGAAGTTATTACTAAGCCTTAGTCGTGGCCCCAGGTAATTTCCTCCCAGGCCTCCATCGGGT   400

                                                                    -30
                                                                    MetAlaGlyProAlaThrGlnSerProM
TATGTATAAAGGGCCCCCTAGAGCTGGGCCCCAAAACAGCCCCGGAGCCTGCAGCCCAGCCCCACCCAGACCCATGGCTGGACCTGCCACCCAGAGCCCCA   500

 -20  -18
etLysLeuMetA
TGAAGCTCATGGGTGAGTGTCTTGGCCCAGGATGGGAGAGCCGCCTGCCCTGGCATGGGAGGGAGGCTGGTGTGACAGAGGGGCTGGGGATCCCCGTTCT   600

                                                                                    -16
                                                                                    laLeuGlnLeuL
GGGAATGGGGATTAAAGGCACCCAGTGTCCCCGAGAGGGCCTCAGGTGGTAGGGAACAGCATGTCTCCTGAGCCCGCTCTGTCCCCAGCCCTGCAGCTGC   700
```

Table VIII (continued)

```
      -10                          -1 +1                                10                                20
euLeuTrpHisSerAlaLeuTrpThrValGlnGluAlaThrProLeuGlyProAlaSerSerLeuProGlnSerPheLeuLeuLysCysLeuGluGlnVa
TGCTGTGGCACAGTGCACTCTGGACAGTGCAGGAAGCCCACCCCCCTGGGCCCTGCCAGCTCCCTGCCCCAGAGCTTCCTGCTCAAGTGCTTAGAGCAAGT   800

                              30             35
lArgLysIleGlnGlyAspGlyAlaAlaLeuGlnGluLysLeu
GAGGAAGATCCAGGGCGATGGCGCAGCGCTCCAGGAGAAGCTGGTGAGTGAGGTGGGTGAGAGGGCTGTGGAGGGAAGCCCCGTGGCGAGAGCTAAGGGG   900

GATGGAACTGCAGGGCCAACATCCTCTGGAAGGGACATGGCAGAATATTAGGAGCAGTGCAGCTGGGGAAGGCTGGGAAGGGACTTGGGGAGGAGGACCT   1000

TGGTGGGGACAGTGCTCGGGAGGGCTGGCTGGGATGGGACTGGAGGCATCACATTCAGGAGAAAGGGCAAGGGCCCCTGTGAGATCAGAGAGTGGGGGTG   1100

CAGGGCAGAGAGGAACTGAACAGCCTGGCAGGACATGGAGGGAGGGGAAAGACCAGAGAGTCGGGGAGGACCCGGGAAGGAGCGGCGACCCGGCCACGGC   1200

                              36          40                          50
                             CysAlaThrTyrLysLeuCysHisProGluGluLeuValLeuLeuGlyHisSerLeuGlyIleProTrpA
GAGTCTCACTCAGCATCCTTCCATCCCCAGTGTGCCACCTACAAGCTGTGCCACCCCGAGGAGCTGGTGCTGCTCGGACACTCTCTGGGCATCCCCTGGG   1300
```

Table VIII (continued)

```
   60                                70 71
laProLeuSerSerCysProSerGlnAlaLeuGlnLeu
CTCCCCTGAGCAGCTGCCCCAGCCAGGCCCTGCAGCTGGTGAGTGTCAGGAAAGGATAAGGCTAATGAGGAGGGGGAAGGAGAGGAGGAACACCCATGGC   1400

                                                                                72
                                                                              AlaGlyCysLeuSerGln
CTCCCCCATGTCTCCAGGTTCCAAGCTGGGGGCCTGACGTATCTCAGGCAGCACCCCCTAACTCTTCCGCTCTGTCTCACAGGCAGGCTGCTTGAGCCAA   1500

      80                                90                                100                               110
LeuHisSerGlyLeuPheLeuTyrGlnGlyLeuLeuGlnAlaLeuGluGlyIleSerProGluLeuGlyProThrLeuAspThrLeuGlnLeuAspValA
CTCCATAGCGGCCTTTTCCTCTACCAGGGGCTCCTGCAGGCCCTGGAAGGGATCTCCCCCGAGTTGGGTCCCACCTTGGACACACTGCAGCTGGACGTCG   1600

                          120
laAspPheAlaThrThrIleTrpGlnGln
CCGACTTTGCCACCACCATCTGGCAGCAGGTGAGCCTTGTTGGGCAGGGTGGCCAAGGTCGTGCTGGCATTCTGGGCACCACAGCCGGGCCTGTGTATGG   1700

                                                                                121
                                                                              MetGluG
GCCCTGTCCATGCTGTCAGCCCCCAGCATTTCCTCATTTGTAATAACGCCCACTCAGAAGGCCCCAACCACTCATCACAGCTTTCCCCCACAGATGGAAG   1800
```

Table VIII (continued)

```
                    130                               140                               150
luLeuGlyMetAlaProAlaLeuGlnProThrGlnGlyAlaMetProAlaPheAlaSerAlaPheGlnArgArgAlaGlyGlyValLeuValAlaSerHi
AACTGGGAATGGCCCCTGCCCTGCAGCCCACCCAGGGTGCCATGCCGGCCTTCGCCTCTGCTTTCCAGCGCCGGGCAGGACGGGTCCTGGTTGCCTCCCA   1900

          160                               170        174
sLeuGlnSerPheLeuGluValSerTyrArgValLeuArgHisLeuAlaGlnProOP
TCTGCAGAGCTTCCTGGAGGTGTCGTACCGCGTTCTACGCCACCTTGCCCAGCCCTGAGCCAAGCCCTCCCCATCCCATGTATTTATCTCTATTTAATAT   2000

TTATGTCTATTTAAGCCTCATATTTAAAGACAGGGAAGAGCAGAACGGAGCCCCAGGCCTCTGTGTCCTTCCCTGCATTTCTGAGTTTCATTCTCCTGCC   2100

TGTAGCAGTGAGAAAAAGCTCCTGTCCTCCCATCCCCTGGACTGGGAGGTAGATAGGTAAATACCAAGTATTTATTACTATGACTGCTCCCCAGCCCTGG   2200

CTCTGCAATGGGCACTGGGATGAGCCGCTGTGAGCCCCTGGTCCTGAGGGTCCCCACCTGGGACCCTTGAGAGTATCAGGTCTCCCACGTGGGAGACAAG   2300

AAATCCCTGTTTAATATTTAAACAGCAGTGTTCCCCATCTGGGTCCTTGCACCCCTCACTCTGGCCTCAGCCGACTGCACAGCGGCCCCTGCATCCCCTT   2400

GGCTGTGAGGCCCCTGGACAAGCAGAGGTGGCCAGAGCTGGGAGGCATGGCCCTGGGGTCCCACGAATTTGCTGGGGAATCTCGTTTTTCTTCTTAAGAC   2500

TTTTGGGACATGGTTTGACTCCCGAACATCACCGACGTGTCTCCTGTTTTTCTGGGTGGCCTCGGGACACCTGCCCTGCCCCCACGAGGGTCAGGACTGT   2600
```

Table VIII (continued)

```
GACTCTTTTTAGGGCCAGGCAGGTGCCTGGACATTTGCCTTGCTGGATGGGGACTGGGGATGTGGGAGGGAGCAGACAGGAGGAATCATGTCAGGCCTGT   2700

GTGTGAAAGGAAGCTCCACTGTCACCCTCCACCTCTTCACCCCCCACTCACCAGTGTCCCCTCCACTGTCACATTGTAACTGAACTTCAGGATAATAAAG   2800

TGTTTGCCTCCAGTCACGTCCTTCCTCCTTCTTGAGTCCAGCTGGTGCCTGGCCAGGGGCTGGGGAGGTGGCTGAAGGGTCGGAGAGGCCAGAGGGAGGT   2900

CGGGGACGAGGTCTGGGGAGGAGGTCCAGGGAGGAGGAGGAAAGTTCTCAAGTTCGTCTGACATTCATTCCGTTAGCACATATTTATCTGAGCACCTACT   3000

CTGTGCAGACGCTGGGCTAAGTGCTGGGGACACAGCAGGGAACAAGGCAGACATGGAATCTGCACTCGAG                                   3070
```

A restriction endonuclease map (approximately 3.4Kb) of the genomic DNA containing the hpG-CSF gene is detailed in Drawing 1. The restriction endonucleases shown in Drawing 1 are: NcoI, N; PstI, P; BamHI, B; ApaI, A; XhoI, X; and Kpn, K. The arrows below the map indicate the sequencing strategy used to obtain the genomic sequence. The boxed regions are those found in the cDNA clone with the dashed box at the end representing the sequence not present in the cDNA clone but identified by probing an mRNA blot. The identification of the coding sequence proposed for exon one was achieved by Northern blot analysis. A 24-mer oligonucleotide probe with the sequence

5' CAGCAGCTGCAGGGCCATCAGCTT3',

spanning the predicted splice junctures for exons 1 and 2 was hybridized to hpG-CSF mRNA in a Northern blot. The blot analysis result shows an mRNA in the same size (~1650 bp) as that observed with an exon 2 oligonucleotide probe. This result and the capability of direct expression of hpG-CSF from the pSVGM-Ppol vector (Example 9) using the Met initiation codon shown in Table VIII, taken together, define the coding sequence contained in exon 1. Exons 2-5 are defined by the coding sequences obtained in the cDNA clone (Ppo2) of the hpG-CSF gene (Table VII).

Practical Example 6

This practical example is related to preparation of an artificial gene coding for hpG-CSF and containing codons preferred for E. coli.

Briefly, the procedures used were generally described in the disclosure of the co-owners Alton et al.'s PCT Application Publication No. WO83/04053, which is contained as a reference herein. The gene was first considered to be a collection of component oligonucleotides assembled into multiple double-stranded chains, which were then assembled into three discrete sections. These sections were designed for ready amplification and, upon removal from the amplification system, could be assembled sequentially or through a multiple fragment ligation in an appropriate expression vector.

The construction of sections I, II, and II is shown in Tables IX-XIV. In the construction of section I, as shown in Tables IX and X, oligonucleotides 1-14 were assembled into 7 double-stranded chains (1 and 8; 2 and 9; 3 and 10; 4 and 11; 5 and 12; 6 and 13; and 7 and 14). The 7 double-stranded chains were then ligated to form section I as shown in Table X. It should also be noted that section I contained an upstream XbaI sticky end and a downstream BamHI sticky end useful for ligation to amplification and expression vectors and for ligation to section II.

Table IX

EChpG-CSF DNA section I

```
CTAGAAAAAACCAAGGAGGTAATAAA      1
TAATGACTCCATTAGGTCCTGCTTCTTCT   2
CTGCCGCAAAGCTTTCTGCTGAAATGTCTGG 3
AACAGGTTCGTAAAATCCAGGGTGACGGT   4
GCTGCACTGCAAGAAAAACTGTGGGCTA    5
CTTACAAACTGTGCCATCCGGAAGAGC     6
TGGTACTGCTGGGTCATTCTCTTGG       7
CATTATTTATTACCTCCTTGGTTTTTT     8
GCAGAGAAGAAGCAGGACCTAATGGAGT    9
TGTTCCAGACATTTCAGCAGAAAGCTTTGCG 10
CAGCACCGTCACCCTGGATTTTACGAACC   11
TAAGTAGCGCACAGTTTTTCTTGCAGTG    12
ACCAGCTCTTCCGGATGGCACAGTTTG     13
GATCCCAAGAGAATGACCCAGCAGT       14
```

Table X

EChpG-CSF DNA section I

```
        10     1    20          30          40    2     50          60
CTAGAAAAA ACCAAGGAGG TAATAAATAA TGACTCCATT AGGTCCTGCT TCTTCTCTGC
    TTTTT TGGTTCCTCC ATTATTTATT ACTGAGGTAA TCCAGGACGA AGAAGAGACG
              8                                        9
XbaI
        70     3    80          90          100   4  110         120
CGCAAAGCTT TCTGCTGAAA TGTCTGGAAC AGGTTCGTAA AATCCAGGGT GACGGTGCTG
GCGTTTCGAA AGACGACTTT ACAGACCTTG TCCAAGCATT TTAGGTCCCA CTGCCACGAC
            10                                       11

       130   5    140         150         160   6   170         180
CACTGCAAGA AAAACTGTGC GCTACTTACA AACTGTGCCA TCCGGAAGAG CTGGTACTGC
GTGACGTTCT TTTTGACACG CGATGAATGT TTGACACGGT AGGCCTTCTC GACCATGACG
            12                                       13

   7   190         100
TGGGTCATTC TCTTGG
ACCCAGTAAG AGAACCCTAG
    14
              BamHI
```

As shown in Tables XI and XII, in the construction of section II, oligonucleotides 15-30 were assembled into 8 double-stranded chains (15 and 23; 16 and 24; 17 and 25; 18 and 26; 19 and 27; 20 and 28; 21 and 29; and 22 and 30). These 8 double-stranded chains were then ligated to form section II, as shown in Table XII. As further shown in Table XII, section II had an upstream BamHI sticky end and a downstream EcoRI sticky end useful for ligation to an amplification vector and for ligation to section I. Near its downstream end, section II also contained a downstream SstI site useful in the eventual ligation of sections II and III.

Table XI

EChpG-CSF DNA section II

GATCCCGTGGGCTCCGCTGTCTTCT	15
TGTCCATCTCAAGCTCTTCAGCTGGC	16
TGGTTGTCTGTCTCAACTGCATTCTGGT	17
CTGTTCCTGTATCAGGGTCTTCTG	18
CAAGCTCTGGAAGGTATCTCTCCGGA	19
ACTGGGTCCGACTCTGGACACTCTGCA	20
GCTAGATGTAGCTGACTTTGCTACTACT	21
ATTTGGCAACAGATGGAAGAGCTCAAAG	22
GACAAGAAGACAGCGGAGCCCACGG	23
ACCAGCCAGCTGAAGAGCTTGAGATG	24
ACAGACCAGAATGCAGTTGAGACAGACA	25
CTTGCAGAAGACCCTGATACAGGA	26
CAGTTCCGGAGAGATACCTTCCAGAG	27
TAGCTGCAGAGTGTCCAGAGTCGGACC	28
AAATAGTAGTAGCAAAGTCAGCTACATC	29
AATTCTTTGAGCTCTTCCATCTGTTGCC	30

Table XII

EChpG-CSF DNA section II

```
      10      15  20          30          40      16  50          60
GATCCCGTG GGCTCCGCTG TCTTCTTGTC CATCTCAAGC TCTTCAGCTG GCTGGTTGTC
    GGCAC CCGAGGCGAC AGAAGAACAG GTAGAGTTCG AGAAGTCGAC CGACCAACAG
               23                                  24
BamHI
      70 17      80          90      18 100          110     19 120
TGTCTCAACT GCATTCTGGT CTGTTCCTGT ATCAGGGTCT TCTGCAAGCT CTGGAAGGTA
ACAGAGTTGA CGTAAGACCA GACAAGGACA TAGTCCCAGA AGACGTTCGA GACCTTCCAT
          25                          26                      27

      130         140   20   150         160         170   21   180
TCTCTCCCGA ACTGGGTCCG ACTCTGGACA CTCTGCAGCT AGATGTAGCT GACTTTGCTA
AGAGAGGCCT TGACCCAGGC TGAGACCTGT GAGACGTCGA TCTACATCGA CTGAAACGAT
                        28                              29

      190         200  22    210
CTACTATTTG GCAACAGATG GAAGAGCTCA AAG
GATGATAAAC CGTTGTCTAC CTTCTCGAGT TTCTTAA
                       30
                          SstI        EcoRI
```

Finally, section III was constructed as shown in Tables XIII and XIV. For this construction, oligonucleotides 31-42 were assembled into 6 double-stranded chains (31 and 37; 32 and 38; 33 and 39; 34 and 40; 35 and 41; and 36 and 42). The 6 double-stranded chains were then ligated to form section III as shown in Table XIV. As also shown in Table XIV, section III

contained an upstream BamHI sticky end and a downstream EcoRI sticky end useful for ligating into an amplification vector, and at least in the case of the EcoRI end, into an expression vector. In addition, section III had an upstream SstI site useful in the eventual ligation of sections II and III.

Table XIII

EChpG-CSF DNA section III

GATCCAAAGAGCTCGGTATGGCACCAG	31
CTCTGCAACCGACTCAAGGTGCTATGCCG	32
GCATTCGCTTCTGCATTCCAGCGTCGTGC	33
AGGAGGTGTACTGGTTGCTTCTCATCTG	34
CAATCTTTCCTGGAAGTATCTTACCGTGT	35
TCTGCGTCATCTGGCTCAGCCGTAATAG	36
AGAGCTGGTGCCATACCGAGCTCTTTG	37
ATGCCGGCATAGCACCTTGAGTCGGTTGC	38
TCCTGCACGACGCTGGAATGCAGAAGCGA	39
ATTGCAGATGAGAAGCAACCAGTACACC	40
CAGAACACGGTAAGATACTTCCAGGAAAG	41
AATTCTATTACGGCTGACCCAGATGACG	42

Table XIV

EChpG-CSF DNA section III

```
        10        31  20          30          40        32  50          60
GATCCAAAG AGCTCGGTAT GGCACCAGCT CTGCAACCCA CTCAAGGTGC TATGCCGGCA
    GTTTC TCGAGCCATA CCGTGGTCGA GACGTTGGGT GAGTTCCACG ATACGGCCGT
                  37                                  38
BamHI       SstI

        70        33  80          90          100       34 110         120
TTCGCTTCTG CATTCCAGCG TCGTGCAGGA GGTGTACTGG TTGCTTCTCA TCTGCAATCT
AAGCGAAGAC GTAAGGTCGC AGCACGTCCT CCACATGACC AACGAAGAGT AGACGTTAGA
                  39                                  40

   35  130         140         150     36  160         170
TTCCTGGAAC TATCTTACCG TGTTCTGCGT CATCTGGCTC AGCCGTAATA G
AAGGACCTTC ATAGAATGGC ACAAGACGCA GTAGACCGAG TCGGCATTAT CTTAA
    41                                      42
                                                      EcoRI
```

The XbaI-BamHI fragment formed by section I was ligated into an M13mp11 phage vector opened with XbaI and BamHI. The vector was then opened again with BamHI and EcoRI digestion, followed by ligation with the BamHI-EcoRI fragment formed by section II. At this point, sections I and II had been joined in a proper orientation. Next, another M13mp11 vector

was opened by BamHI and EcoRI digestion and then ligated with the BamHI-EcoRI fragment formed by section III.

The vector containing section I and II was digested with XbaI and SstI. Likewise, the vector containing section III was digested with SstI and EcoRI. Both of the smaller of the two fragments resulting from each digestion were ligated into a plasmid pCFM1156 which had been opened with XbaI and EcoRI. The product of this reaction was an expression plasmid containing a continuous DNA sequence, as shown in Table XV, it s DNA sequence is encoding the entire hpG-CSF polypeptide with an amino terminal methionine codon (ATG) for E. coli translation initiation.

Table XV

```
                                          -1  +1
                                         Met Thr Pro Leu Gly Pro Ala Ser Ser Leu
  C TAG AAA AAA CCA ACG AGC TAA TAA ATA ATG ACT CCA TTA GGT CCT GCT TCT TCT CTG

 10                                       20
CCG CAA AGC TTT CTG CTG AAA TGT CTG GAA CAG GTT CGT AAA ATC CAG GGT GAC GGT GCT
Pro Gln Ser Phe Leu Leu Lys Cys Leu Glu Gln Val Arg Lys Ile Gln Gly Asp Gly Ala

 30                                       40
Ala Leu Gln Glu Lys Leu Cys Ala Thr Tyr Lys Leu Cys His Pro Glu Glu Leu Val Leu
GCA CTG CAA GAA AAA CTG TGC GCT ACT TAC AAA CTG TGC CAT CCG GAA GAG CTG GTA CTG

 50                                       60
Leu Gly His Ser Leu Gly Ile Pro Trp ala Pro Leu Ser Ser Cys Pro Ser Gln Ala Leu
CTG GGT CAT TCT CTT GGC ATC CCG TGG GCT CCG CTG TCT TCT TGT CCA TCT CAA GCT CTT

 70                                       80
Gln Leu Ala Gly Cys Leu Ser Gln leu His Ser Gly Leu Phe Leu Tyr Gln Gly Leu Leu
CAG CTG GCT GGT TGT CTG TCT CAA CTG CAT TCT GGT CTG TTC CTG TAT CAG GGT CTT CTG

 90                                       100
Gln Ala Leu Glu Gly Ile Ser Pro Glu Leu Gly Pro Thr Leu Asp Thr Leu Gln Leu Asp
CAA GCT CTG GAA GGT ATC TCT CCG GAA CTG GGT CCG ACT CTG GAC ACT CTG CAG CTA GAT

110                                      120
Val Ala Asp Phe Ala Thr Thr Ile Trp Gln Gln Met Glu Glu leu Gly Met Ala Pro Ala
GTA GCT GAC TTT GCT ACT ACT ATT TGG CAA CAG ATG GAA GAG CTC GGT ATG GCA CCA GCT

130                                      140
Leu Gln Pro Thr Gln Gly Ala Met Pro Ala Phe Ala Ser Ala Phe Gln Arg Arg Ala Gly
CTG CAA CCG ACT CAA GGT GCT ATG CCG GCA TTC GCT TCT GCA TTC CAG CGT CGT GCA GGA
```

Table XV (cont'd)

```
150                                      160
Gly Val Leu Val Ala Ser His Leu Gln Ser Phe Leu Glu Val Ser Tyr Arg Val Leu Arg
GGT GTA CTG GTT GCT TCT CAT CTG CAA TCT TTC CTG GAA GTA TCT TAC CGT GTT CTG CGT

170              174
His Leu Ala Gln Pro
CAT CTG GCT CAG CCG TAA TAG AAT T
```

Although any appropriate vector could be used to express this DNA, the expression plasmid pCFM1156 could be easily constructed from pCFM836, the construction of which was described in published European Patent Application No. 136,490. pCFM836 was first cut with NdeI and then blunt-ended with PolI so that both the NdeI sites were destroyed. Next, the vector was digested with ClaI and SacII to remove a pre-existing polylinker before ligation to a substitute polylinker as shown in Table XVI. This substitute polylinker could be constructed according to the procedures of Alton et al. (ibid.). Control of expression in the expression plasmid pCFM1156 was by a lambda P_L promoter, which itself could be under the control of a CI857 repressor gene (as can be obtained from E. coli strain K12ΔHtrp).

Table XVI

```
1   ATCGATTTGATTCTAGAAGGAGGAATAACATATGGTTAACGCGTTCGAATTCCGTACCAT
    TAGCTAAACTAAGATCTTCCTCCTTATTGTATACCAATTGCGCAACCTTAAGCCATGGTA

    1 ClaI, 12 XbaI, 29 NdeI, 35 HincII, HpaI, 39 MluI, 47 EcoRII,
      53 HgiCI KpnI, 57 NcoI StyI,

61  GGAAGCTTACTCGAGGATCCGCGGATAAATAAGTAACGATCC
    CCTTCGAATGAGCTCCTAGGCGCCTATTTATTCATTGCTAGG

    63 HindIII, 70 AvaI XhoI, 75 BamHI Xho2, 79 Sac2,
```

Practical Example 7

This practical example is related to E. coli expression of an hpG-CSF polypeptide by means of a DNA sequence encoding [Met-^{1}] hpCSF. The sequence used was partially synthetic and partially cDNA-derived. The synthetic sequences used were codons preferred for E. coli.

Plasmid Ppo2, containing the hpG-CSF gene shown in Table VII, was digested with HgiAI and SsuI, thereby obtaining an approximately 645 base pair fragment including the gene for mature hpCSF (as shown in Table VII) with seven of the leader sequence residue codons at the 5' end and about 100 base pairs of the 3' non-coding region. HgiAI digestion left a 4-base sticky end at 5', identical to that of PstI digestion, and StuI left a blunt end. This allowed for ready insertion of the fragment into M13mp8 (Rf) cut with PstI and with the blunt-end-forming restriction enzyme HindII. Upon amplification in M13, the hpG-CSF DNA was excised by digestion with ApaI and BamHI which cut, respectively, at a ApaI site spanning the codons for residues +3 to +5 of hp-CSF and at a BamHI site "downstream" of the HindII site. To make the expression of hpG-CSF polypeptide in E. coli possible, a synthetic fragment was prepared as shown in Table XVII.

Table XVII

```
5' - C.TAG AAA AAA CCA AGG AGG TAA TAA ATA
      3' - TTT TTT GGT TCC TCC ATT ATT TAT
     XbaI

-1  +1
Met Thr Pro Leu
ATG ACA CCT CTG GGC C - 5'
TAC TGT GGA GAC -3'
                ApaI
```

As can be determined from analysis of Table XVII, the linker included an ApaI sticky end, codons designating the initial three residues of the amino terminal of hpG-CSF ("restoring" the Thr^1, Pro^2, Leu^3-designating codons deleted upon ApaI digestion of the M13 DNA described above and using codons preferentially expressed in E. coli), a translation initiating ATG, a sequence of 24 base pairs providing a ribosome binding site, and an XbaI sticky end.

The expression vector used for E. coli expression was one described as pCFM536 in European Patent Application No. 136,490, by Morris, published on April 10, 1985. Also see A.T.C.C. 39934 and E. coli strain JM103 containing pCFM536. Briefly, plasmid pCFM536 was digested with XbaI and BamHI. The hpG-CSF fragment (ApaI/BamHI) and linker (XbaI/APAI) described above were then ligated there to form a plasmid designated p536Ppo2.

The plasmid p536Ppo2 was transformed into a phage resistant variant of E. coli strain AM7, which had been transformed with plasmid pMW1 (A.T.C.C. No. 39933) containing a CI^{857} gene. Transformation was verified on the basis of the antibiotic (amp) resistance marker gene carried on the pCFM536 progenitor plasmid. The cells were cultured in LB broth (ampicillin at 50 µg/mL) at 28°C. Upon growth of the cells in culture to A_{600} = 0.5, hpCSF expression was induced by increasing the culture temperature to 42°C for 3 hr. The final O.D. of the culture was A_{600} = 1.2.

The level of expression of hpG-CSF by the transformed cells was estimated on SDS-polyacrylamide gel stained with coomassie blue to be 3-5% of total cellular proteins.

The cells were harvested by centrifugation at 3500 g for 10 min in a JS-4.2 rotor. The cells at 25% (w/v) in water were broken by passing three times through a French Pressure Cell at 10,000 psi (700 kg/cm^2). The suspension of broken cells was centrifuged at 10,000 x g for 15 min in a JA-20 rotor. The pellet was resuspended in water and solubilized at about 5 mg/mL total protein in 1% lauric acid, 50 mM Tris, pH 8.7. The solubilized pellet material was centrifuged at 15,000 x g for 10 min, and $CuSO_4$ was added to 20 mM to the supernatant. After

1 hr, this sample was applied to a C4 HPLC column for purification according to the procedures of Practical Example 1 (B) with modifications made for volume and concentration.

A second purification procedure was developed to produce larger quantities of hpG-CSF prepared in an organic solvent-free buffer. This material is suitable for in vivo studies. 150 g of a cell paste was resuspended in about 600 mL of 1 mM DTT at about 7000 psi (490 kg/cm^2). The suspension of broken cells was centrifuged at 10,000 x g for 30 min, and the pellet was resuspended in 400 mL of 1% deoxycholate (DOC), 5 mM EDTA, 5 mM DTT, and 50 mM Tris, pH 9. This suspension was stirred at room temperature for 30 min, and then centrifuged at 10,000 x g for 30 min. The pellet was resuspended in about 400 mL of water and centrifuged at 10,000 g for 30 min. The pellet was solubilized in 100 mL of 2% Sarkosyl and 50 mM Tris at pH 8. $CuSO_4$ was added to 20 μM and the mixture was stirred for 16 hr at room temperature, followed by centrifugation at 20,000 x g for 30 min. Then, 300 mL of acetone was added to the supernatant. This mixture was allowed to stand on ice for 20 min and then centrifuged at 5000 x g for 30 min. The pellet was dissolved in 250 mL of 6 M guanidine and 40 mM sodium acetate at pH 4, and applied to a 1,200 mL G-25 column equilibrated and run in 20 mM sodium acetate at pH 5.4. The hpG-CSF peak (about 400 mL) was pooled, and applied onto a 15 mL CM-cellulose column equilibrated in 20 mM sodium acetate at pH 5.4. After loading, the column was washed with 60 mL of 20 mM sodium acetate at pH 5.4 and 25 mM sodium chloride, and then the column was eluted with 200 mL of 20 mM sodium acetate at pH 5.4 and 37 mM sodium chloride. 150 mL of this eluent was concentrated to 10 mL and applied to a 300 mL G-75 column equilibrated and run in 20 mM sodium acetate and 100 mM sodium chloride at pH 5.4. The peak fractions in 35 mL were pooled and sterilized by filtration. The final concentration of hpG-CSF was 1.5 mg/mL, with a greater than 95% purity as determined by analysis on a gel, which contained less than 0.5 ng of pyrogen per 0.5 mg of hpG-CSG. The pyrogen level was determined using a Limulus Amebocyte Lysate (LAL) test kit (M. A. Bioproducts, Walkersville, Maryland).

Practical Example 8

This practical example is related to use of recombinant methods to generate analogs of hpG-CSF with each of cysteine residues present at positions 17, 36, 42, 64 and 74 replaced by an appropriate amino acid residue.

Site-directed mutagenesis procedures according to Souza et al. in PCT Application No. WO85/00817 published on February 28, 1985 were carried out on [Met^{-1}] encoding DNA of plasmid p536Ppo2 described below, using synthetic oligonucleotides ranging in size from 20 to 23 bases shown in Table XVIII. Oligonucleotide No. 1 allowed for formation of a gene encoding [Ser^{17}]hpG-CSF while oligonucleotide No. 2 allowed for formation of [Ser^{36}]hpG-CSF.

Table XVIII

Oligonucleotide	Sequence
1.	5'-CTG CTC AAG TCC TTA GAG CAA GT-3'
2.	3'-GAC AAG CTG TCT GCC ACC TACA-3'
3.	5'-TAC AAG CTG TCC CAC CCC GAG-3'
4.	5'-TGA GCA GCT CCC CCA GCC AG-3'
5.	5'-CTG GCA GGC TCC TTG AGC CAA-3'

The restrictive Cys to Ser site-directed mutagenesis was carried out using M13 mp10 containing an XbaI-BamHI hpG-CSF fragment isolated from p536Ppo2 as a template. DNA from each M13mp13 clone containing a Cys-Ser substitution was treated with XbaI and BamHI. The resulting fragment was cloned into expression vector pCFM746 and the expression product was isolated as in Practical Example 7.

The plasmid pCFM746 could be constructed by cleaving the plasmid pCFM736 (the construction of which from deposited and publicly available materials is described by Morris in PCT Application No. WO85/00829, published on February 28, 1985) with ClaI and BamHI to remove a pre-existing polylinker and by substituting the following polylinker.

Table XIX

```
   ClaI
5'CGATTTGATTCTAGAATTCGTTAACGGTACCATGGAA
3'  TAAACTAAGATCTTAAGCAATTGCCATGGTACCTT

GCTTACTCGAGGATCCGCGGATAAATAAGTAAC3'
CGAATGAGCTCCTAGGCGCCTATTTATTCATTGCTAG5'
                                  Sau3a
```

In a purification procedure for the Cys-Ser (from Cys to Ser) analog according to the present invention, about 10-15 g of a cell paste was resuspended in 40 mL of 1 mM DTT and

passed 3 times through a French Pressure Cell at 10,000 psi (700 kg/cm^2). The suspension of broken cells was centrifuged at 1,000 x g for 30 min. The pellet was resuspended in 1% DOC, 5 mM EDTA, 5 mM DTT, 50 mM Tris, pH 9 and was stirred for 30 min at room temperature. The mixture was centrifuged at 10,000 x g for 30 min. The pellet was resuspended in 40 mL H_2O, and then re-centrifuged at 10,000 x g for 30 min. The pellet was dissolved in 10 mL of 2% Sarkosyl, 50 mM DTT, 50 mM Tris, pH 8. After stirring for 1 hr, the mixture was clarified by centrifugation at 20,000 x g for 30 min, and then applied to a 300 mL G-75 column equilibrated and run in 1% Sarkosyl, 50 mM Tris, pH 8. Fractions containing the analog were pooled and allowed to oxidize by air by standing with exposure to air for at least one day. The final concentration ranged from 0.5 to 5 mg/mL.

Practical Example 9

In this practical example, a mammalian cell expression system was developed to ascertain expression of an active polypeptide product of hpG-CSF DNA in and secretion from mammalian cells (COS-1, A.T.C.C. CRL-1650). This system was designed to provide for secretion of a polypeptide analog of hpG-CSF via expression and secretory processing of a partially synthetic, partially cDNA-derived construct encoding [Ala^1] hpG-CSF preceded by a leader polypeptide having the sequence of residues attributed to human GM-CSF described by Wong et al. [Science, 228, 810-815 (1985)] and by Lee et al. [Proc. Natl. Acad. Sci. (USA), 82, 4360-4364 (1985)].

The expression vector used for preliminary studies of expression of a polypeptide product of the present invention was a "shuttle" vector containing both pBR322 and SV40 DNA, which had been designed to allow for autonomous replication in both E. coli and mammalian cells, with mammalian cell expression of an inserted exogenous DNA under control of a viral promoter/regulator DNA sequence. This vector, designated pSVDM-19, harbored in E. coli HB101, was deposited on August 23, 1985 with the American Type Culture Collection (12301 Parklawn Drive, Rockville, Maryland), and received the accession No. A.T.C.C. 53241.

The specific manipulations involved in the expression vector construction were as follows. A leader-encoding DNA sequence was synthesized as shown in Table XX.

Table XX

```
                                        -17
                HindIII                 Met Trp
          5' - A GCT TCC AAC ACC ATG TGG
               3' - AGG TTG TGG TAC ACC

                                -10
Leu Gln Ser Leu Leu Leu Leu Gly Thr Val
CTG CAG AGC CTG CTG CTC TTG GGC ACT GTG
GAC GTC TCG GAC GAC GAG AAC CCG TGA CAC

                        -1  +1
Ala Cys Ser Ile Ser Ala Pro Leu
GCC TGC AGC ATC TCT GCA CCC CTG GGC G -3'
CGG ACG TCG TAG AGA CGT GGG GAC -5'
                                        ApaI
```

As shown in Table XX, the sequence contained HindIII and ApaI sticky ends and codons for the 17 amino acid residues attributed to the "leader" of human hpG-CSF. They were followed by codons specifying an alanine residue, a proline residue and a leucine residue. The proline and leucine residues duplicated the amino acids present at positions +2 and +3 of hpG-CSF, while the alanine residue was duplicative of the initial amino terminal (+1) residue of GM-CSF rather than hpG-CSF. Substitution of the threonine by alanine was designed to facilitate the "processing off" of the GM-CSF leader by cellular mechanisms ordinarily involved in GM-CSF secretion processing in an appropriate host cell.

The plasmid pSVDM-19 was digested with KpnI and the site was converted to a blunt end with Klenow enzyme. Then, the DNA was cut with HindIII. The resulting large fragment was combined and ligated with the HindIII/PvuIII fragment shown in Table VII (isolated from plasmid Ppo2 as the second largest fragment resulting from HindIII digestion and partial digestion with PvuII) to obtain plasmid pSV-Ppo1. The artificial GM-CSF leader sequence fragment of Table VIII was then ligated into pSV-Ppo1 (following its cleavage with HindIII and ApaI) to obtain plasmid pSVGM-Ppo1.

Calcium phosphate precipitate (1-5 μg) of plasmid pSVGM-Ppo1 DNA was transformed into duplicate 60 mm plates of COS-1 cells essentially as described by Wigler et al. [Cell, 14, 725-731 (1978)]. As a control, plasmid pSVDM-19 was also transformed into COS-1 cells. The culture supernatant of the cells was harvested 5 days after transfection and assayed for hpG-CSF activity. Yields of [Ala^{1}]hpG-CSF from the culture supernatant were of the order of 1 to 2.5 μg/mL.

Following successful expression of the [Ala^1]hpG-CSF product-encoding plasmid pSVGM-Ppo1 in COS-1 cells, another vector was constructed, which contained the human GM-CSF leader sequence but had a codon for a threonine residue (naturally occurring at position 1 of hpG-CSF) replacing the codon for alanine at that position. Briefly, an oligonucleotide was synthesized ($^{5'}$CAGCATCTCTACACCTCTGGG) for site-directed mutagenesis (SDM). The HindIII-BamHI fragment of hpG-CSF in pSVGM-Ppo1 was ligated into M13mp10 for the SDM. The newly synthesized hpG-CSF gene containing a Thr codon in position one was isolated by cleavage with HindIII and EcoRI. The fragment was then cloned into pSVDM-19, prepared by cleavage with the two same restriction endonucleases. The resulting vector pSVGM-Ppo(Thr) was transformed into COS cells and the yields of hpG-CSF measured in the culture supernatant ranged from 1 to 5 μg/mL.

Finally, the genomic sequence whose isolation was described in Example 5 was used to construct an expression vector for mammalian cell expression of hpG-CSF. More specifically, pSVDM-19 was digested with KpnI and HindIII and the large fragment used in a four-way ligation with a synthetic linker with HindIII and NcoI sticky ends, as shown in Table XXI. Exxon 1 containing NcoI-BamHI fragment was isolated from pBR322 (8500 hpG-CSF genomic subclone) and exons 2-5 containing BamHI-Kpn1 fragment were isolated from the plasmid pBR322 (8500 hpG-CSF genomic subclone). The resulting mammalian expression vector, pSV/ghG-CSF, produced 1 to 2.5 μg/mL of hpG-CSF from transformed COS cells.

Table XXI

HindIII
5'AGCTTCCAACAC
AGGTTGTGGTAC5'
NcoI

Example 10

This practical example is related to the physical and biological properties of the recombinant polypeptide products of the present invention.

1. Molecular Weight

The recombinant hpG-CSF products of E. coli expression as that in Practical Example 7 had an apparent molecular weight of 18.8 kD as determined by reducing SDS-PAGE (as would be predicted from analysis of the deduced amino acid sequence of Table VII), whereas natural isolates purified as described in Practical Example 1 had an apparent molecular weight of 19.6 kD. Presence of N-glycans in the natural isolates could actually be ruled out on the basis of the lack of asparagine residues in the primary sequence of hpG-CSF shown in Table VII. Accordingly, if O-glycans were responsible for the molecular weight difference between natural isolates and the non-glycosylated recombinant products was determined as follows. Approximately 5 μg of the natural isolate material was treated with neuraminidase (Calbiochem, La Jolla, California). A 0.5 μg aliquot was removed, and the remaining material was incubated with 4 mU O-glycanase (enco-x-n-acetylgalactose aminidase, Genzyme, Boston, Massachusetts) at 37°C. Aliquots were removed after ½, 2 and 4 hr of incubation. These samples were subjected to SDS-PAGE side by side with the E. coli derived recombinant material. After neuraminidase treatment, the apparent molecular weight of the isolate shifted from 19.6 kD to 19.2 kD, suggesting removal of a sialic acid residue. After 2 hr of treatment with O-glycanase, the molecular weight shifted to 18.8 kD, which was identical to the apparent molecular weight of the E. coli derived material. The sensitivity of the carbohydrate structure to neuraminidase and O-glycanase suggests the following structure for the carbohydrate component: N-acetylneuraminic acid-α-(2-6)(galactose-ß-(1-3) N-acetylgalactoseamine-R, wherein R is serine or threonine).

2. ^{3}H-Thymidine Incorporation

Induction of proliferation of human bone marrow cells was assayed based on increased incorporation of ^{3}H-thymidine. Human bone marrow from healthy donors was subjected to a density gradient centrifugation with Ficoll-Hypaque (1.077 g/mL, Pharmacia). Low density cells were suspended in Iscove's medium (GIBCO) containing 10% fetal bovine serum and glutamine-penicillin-streptomycin. Subsequently, 2 x 10^4 human bone marrow cells were incubated with either control medium or the recombinant E. coli material of Practical Example 7 in 96-well flat bottom plates at 37°C with 5% CO_2 in air for 2 days. The samples were assayed in duplicate and the concentration varied over a 10,000-fold range. Cultures were then pulsed for 4 hr with 0.5 μ Ci/well of ^{3}H-Thymidine (New England Nuclear, Boston, Massachusetts). The ^{3}H-thymidine incorporation was measured as described by Ventua et al. [Blood, 61, 781 (1983)]. In this assay, human hpG-CSF isolates can induce ^{3}H-thymidine incorporation into human bone marrow cells at a level approximately 4-10 times higher than control supernatant. The E. coli-derived hpG-CSF material of Practical Example 6 showed similar properties.

A second human bone marrow cell proliferation study was carried out using a culture medium of transfected COS-1 cells as prepared in Practical Example 9 and yielded similar results, indicating that the encoded polypeptide products were indeed secreted into the culture medium as active materials.

3. WEHI-3B D^+ Differentiation Induction

Activity of the recombinant, E. coli-derived material to induce differentiation of the murine myelomonocytic leukemic cell line WEHI-3B D^+ was assayed in a semi-solid agar medium as described by Metcalf [Int. J. Cancer, 25, 225 (1980)]. The recombinant hpG-CSF product and medium control were incubated with ~60 WEHI-3B D^+ cells/well at 30°C with 5% CO_2 in air for 7 days. The samples were incubated in 24-well flat bottom plates and the concentration varied over a 2000-fold range. Colonies were classified as undifferentiated, partially differentiated or wholly differentiated and colony cell counts were measured microscopically. The E. coli recombinant material was found to induce differentiation.

4. CFU-GM, BFU-E and CFU-GEMM Assays

Natural isolates of pluripotent human G-CSF (hpG-CSF) and the recombinant pluripotent human G-CSF (hpG-CSF) were found to cause human bone marrow cells to proliferate and differentiate. These activities were measured in CFU-GM [Broxmeyer et al., Exp. Hematol., 5, 87, (1971)], BFU-E and CFU-GEMM assays [Lu et al., Blood, 61, 250 (1983)] using low density, non-adherent bone marrow cells from healthy human volunteers. A comparison of CFU-GM, BFU-E and CFU-GEMM biological activities using either 500 units of hpG-CSF or rhpG-CSF are shown in Table XXII.

All the colony assays were performed with low density non-adherent bone marrow cells. Human bone marrow cells were subject to a density gradient centrifugation with Ficoll-Hypaque (density, 1.077 g/cm^3, Pharmacia). The low density cells were then resuspended in Iscove's modified Dulbecco medium containing fetal calf serum and placed for adherence on a Falcon tissue culture dish (No. 3003, Becton Dickenson, Cockeysville, MD) for 1-1/2 hr at 37°C.

Table XXII

	CFU-GM	BFU-E	CFU-GEMM
Medium	0 ± 0	26 ± 1	0 ± 0
Natural			
hpG-CSF	83 ± 5.4	83 ± 6.7	4 ± 0
rhpG-CSF	87 ± 5	81 ± 0.1	6 ± 2

Medium control consisted of Iscove's modified Dulbecco medium plus 10% FCS, 0.2 mM hemin and 1 unit of recombinant erythropoietin.

For the CFU-GM assay, target cells were plated at 1 x 10^5 in 1 mL of 0.3% agar culture medium containing supplemented McCoy's 5A medium and 10% heat inactivated fetal calf serum. Cultures were scored for colonies (greater than 40 cells per aggregate) and morphology was assessed on day 7 of the culture. The number of colonies is shown as the mean ± SEM as determined from quadruplicate plates.

In the case of BFU-E and CFU-GEMM analyses, cells (1 x 10^5) were added to 1 mL of a mixture containing Iscove's modified Dulbecco's culture medium (Gibco), 0.8% methylcellulose, 30% bovine fetal serum, 0.05 nM 2-mercaptoethanol, 0.2 mM hemin and 1 unit of recombinant erythropoietin and the dish was incubated in 5% CO_2 and 5% O_2 in a humidified atmosphere. The low oxygen pressure was achieved using an oxygen reducer (Reming Bioinstruments, Syracuse, N.Y.). After 14 days of incubation, the colonies were measured. The number of colonies was determined on two identical plates and expressed as the mean ± SEM.

Colonies formed in the CFU-GM assay were all found to be chloracetate esterase positive and non-specific esterase (alpha-naphthyl acetate esterase) negative, consistent with the colonies being granulocyte in type. Both natural hpG-CSF and rhpG-CSF were found to have a specific activity of approximately 1 x 10^8 U/mg pure protein, when assayed by serial dilution in a CFU-GM assay. The BFU-E and CFU-GEMM data in Table XXII are representative of three separate experiments and similar to the data reported previously for natural hpG-CSF. It is important to note that the rhpG-CSF is extremely pure and free of other potential mammalian growth factors by virtue of its production in E. coli. Thus rhpG-CSF is capable of supporting mixed colony formation (CFU-GEMM) and BFU-E when added in the presence of recombinant erythropoietin.

5. Cell Binding Assays

It was previously reported that WEHI-3B(D^+) cells and human leukemic cells from newly diagnosed leukemia patients would bind ^{125}I-labeled mouse G-CSF and that this binding could be completed by addition of unlabeled G-CSF or human CSF-β. The ability of natural hpG-CSF and rhpG-CSF to compete for binding of ^{125}I-hpG-CSF to human and mouse leukemic cells was tested. Highly purified natural hpG-CSF (>95% pure; 1 μg) was iodinated [Tejedor et al., Anal. Biochem., 127, 143 (1982)], and separated from reactants by gel filtration and ion exchange chromatography. The specific activity of the natural ^{125}I-hpG-CSF was approximately 100 μCi/μg protein. Mouse WEHI-3B(D^+) and two human peripheral blood myeloid leukemic cell preparations (ANLL, one classified as M4, and the other as M5B) were tested for their ability to bind ^{125}I-hpG-CSF.

The mouse and freshly prepared human peripheral blood myeloid leukemic cells were washed three times with PBS/1% BSA. WEHI-3B(D^+) cells (5 x 10^6) or fresh leukemic cells (3 x 10^6) were incubated in duplicate in PBS/1% BSA (100 μL) in the absence or presence of various concentrations (volume: 10 μL) of unlabeled hpG-CSF, rhpG-CSF or GM-CSF and in the presence of ^{125}I-hpG-CSF (approx. 100,000 cpm or 1 ng) at 0°C for 90 min (total volume: 120 μL). Cells were then resuspended and layered over 200 μL of ice cold FCS in a 350 μL plastic centrifuge tube and centrifuged (1000 x g; 1 min). The pellet was collected by cutting off the end of the tube and the pellet and supernatant were counted separately in a gamma counter (Packard).

Specific binding (cpm) was determined as total binding in the absence of a competitor (mean of duplicates) minus binding (cpm) in the presence of a 100-fold excess of unlabeled hpG-CSF (non-specific binding). The non-specific binding was maximally 2503 cpm for WEHI-3B(D^+) cells, 1072 cpm for ANLL (M4) cells and 1125 cpm for ANLL (M5B) cells. Experiments 1 and 2 were run on separate days using the same preparation of ^{125}I-hpG-CSF and displayed internal consistency in the percent inhibition observed with 2000 units of hpG-CSF. The results are shown in Table XXIII.

Table XXIII

Competitor		WEH1-3B (D*)		ANLL (M4)		ANLL (M5B)	
Experiment 1	(U/mL)	cpm	Inhibition rate %	cpm	Inhibition rate %	cpm	Inhibition rate %
No	0	6,608	--	1,218	--	122	--
hpG • CSF:	10,000	685	90				
	2,000	1,692	74	94	97	-376	0
	200	2,031	69				
rhpG • CSF:	10,000	0	100				
	2,000	1,185	82	202	83	0	0
Experiment 2							
No	0	2,910	0				
Natural hpG • CSF:	2,000	628	78				
GM • CSF:	2,000	3,311	0				

As shown in Table XXIII, ^{125}I-hpG-CSF demonstrated binding to the WEHI-3B(D^+) leukemic cells. The binding was inhibited in a dose dependent manner by unlabeled natural hpG-CSF or rhpG-CSF, but not by GM-CSF. In addition, binding of natural hpG-CSF to human myelomonocytic leukemic cells (ANLL, M4) was observed. The binding to these cells was comparable to the response to natural hpG-CSF in liquid cultures by differentiation into mature macrophages as judged by morphology. The absence of binding of natural ^{125}I-hpG-CSF to monocytic leukemic cells from another patient (ANLL, M5B) suggests that certain leukemias may differentially express or lack receptors for hpG-CSF. The ability of rhpG-CSF to compete for the binding of natural ^{125}I-hpG-CSF, similarly to natural hpG-CSF, suggests that the receptors recognize both forms equally well.

These studies demonstrating the binding of natural ^{125}I-labeled hpG-CSF to leukemic cells were comparable with the ability of natural hpG-CSF in culture to induce granulocytic and monocytic differentiation of low density bone marrow cells obtained from one patient with acute promyelocytic leukemia (M3) and a second patient with acute myeloblastic leukemia (M2). Cells from each patient were cultured for four days in medium alone or in the presence of 1×10^5 unit of hpG-CSF. Cells from the M3 control culture incubated in medium alone were still promyelocyte in type, while cells cultured in the presence of rhpG-CSF appeared to be mature

cells of the myeloid type including metamyelomonocytes. The actual differential counts for this patient, on 100 cells evaluated for the control, there were 100% promyelocytes, but for the rhpG-CSF treated cells, there were 22% blasts plus promyelocytes, 7% myelocytes, 35% metamyelocytes, 20% band forms plus segmented neutrophils, 14% monocytes and 2% macrophages. It is worth noting that one of the polymorphonuclear granulocytes still contained a prominent Auer rod, suggesting that at least this cell represented a differentiated cell belonging to the leukemic clone. Cells from the second patient with myeloblastic leukemia (M2) were also cultured for four days in the presence of absence of rhpG-CSF. Visual analysis of M2 cells cultured in medium alone revealed large "blast-like" cells, some of which had nucleoli. Some of the M2 cells, when treated with rhpG-CSF, differentiated into mature segmented neutrophils displaying residual Auer rods in the center, suggesting differentiation occurring in cells belonging to the leukemic clone. The actual differential counts of 100 cells evaluated morphologically revealed that the control cells consisted of 100% blasts whereas the rhpG-CSF-treated cells consisted of 43% blasts, 1% myelocytes, 15% metamyelocytes, 28% band forms plus segmented neutrophils, 2% promonocytes and 11% monocytes. The leukemic cells were also examined for differentiation at four other concentrations of rhpG-CSF (5×10^3, 1×10^4, 2.5×10^4 and 5×10^4 U/mL, data not shown). Even at the lowest concentration of rhpG-CSF tested (5×10^3 U/mL), there was still significant differentiation (cells differentiated beyond myelocytes) with the M3 (50%) and M2 (37%) leukemic cells.

6. Immunoassays

To prepare polyclonal antibodies for immunoassay use, the antigen used was the pluripotent bladder cancer cell line 5637 G-CSF (1A6) as prepared in Practical Example 1 (B). This material was estimated to be 85% pure based on silver nitrate staining of polyacrylamide gels. Six week-old Balb/C mice were immunized with multiple-site subcutaneous injections of the antigen. The antigen was resuspended in PBS and emulsified with an equal volume of Freund's complete adjuvant. The dose was 5 to 7 μg of antigen per mouse per injection. A booster immunization was administered 18 days later with the same amount of antigen emulsified with an equal volume of Freund's incomplete adjuvant. 4 days later mouse serum was taken to test for the antibody specific for human pluripotent G-CSF.

Dynatech Immulon II Removawell strips in holders (Dynatech Lab., Inc., Alexandria, Virginia) were coated with hpG-CSF at 5 μg/mL in 50 mM carbonate-bicarbonate buffer, pH 9.2. Wells were coated with 0.25 μg in a volume of 50 μL. The antigen-coated plates were incubated for 2 hours at room temperature and overnight at 4°C. The solution was decanted and the plates were incubated for 30 minutes with PBS containing 5% BSA to block the reactive

surface. This solution was decanted and diluted preimmune or test serum were added to the wells and incubated for 2 hr at room temperature. The sera were diluted with PBS, pH 7.0 containing 1% BSA. The serum solution was decanted and the plates were washed three times with Washing Solution (KPL, Gaithersburg, Maryland). Approximately 200,000 cpm of iodinated rabbit anti-mouse IgG (NEN, Boston, Massachusetts) in 50 µL of PBS, pH 7.0, containing 1% BSA was added to each well. After incubation for 1½ hr at room temperature, the solution was decanted and the plates were washed 5 times with Washing Solution. The wells were removed from holder and counted in a Beckman 5500 gamma counter. High-titered mouse sera showed a greater than 12-fold higher reactivity than the corresponding preimmune serum at a dilution of 1:100.

The immunological property of the E. coli-derived hpG-CSF was determined by its reactivity to the high-titered mouse serum specific for mammalian-cell-derived hpG-CSF. Thus, 0.25 µg of 90% pure E. coli-derived protein was coated to Immulon II Removable wells in a volume of 50 µL and the mouse serum was assayed as described above.

The high-titered mouse serum showed a 24-fold higher reactivity to the E. coli-derived material than did the corresponding preimmune serum at a dilution of 1:100.

7. Serine Analog Bioassays

The [Ser^{17}]hpG-CSF, [Ser^{36}]hpG-CSF, [Ser^{42}]hpG-CSF, [Ser^{64}]hpG-CSF, and [Ser^{74}]hpG-CSF products prepared according to Practical Example 9 were assayed for hpG-CSF activity in the ^{3}H-thymidine incorporation, CFU-GM, and WEHI3B D^+ assays. In each assay, the [Ser^{17}] analog had an activity comparable to that of the recombinant molecule having the native structure. The remaining analogs had an activity 100 times lower in the ^{3}H-thymidine incorporation assay, 250 times lower in the CFU-GM assay, and 500 times lower in the WEHI-3B D^+ assay. These results support the contention that the cysteines at positions 36, 42, 64 and 74 may be needed for the full biological activity.

8. In vivo Bioassay

Alzet® osmotic pumps (Alzet Corp., Palo Alto, CA; Model 2001) were connected to in-dwelling right jugular vein catheters and implanted subcutaneously in seven male Syrian golden hamsters. Four of the pumps contained a buffer [20 mM sodium acetate (pH 5.4) and 37 mM sodium chloride] and 1.5 ng/mL E. coli-derived hpG-CSF, while 3 contained the buffer alone. The pumping rate for the osmotic pumps was 1 µL/hr for up to seven days. At the third day after implantation of the pumps, the mean granulocyte count of the four treated hamsters was

six times higher than that of the three (buffer) controls and the increased granulocyte count was reflected in a four-fold increase in total lymphocytes. The erythrocyte count was unchanged by the treatment. These results indicate that the recombinant material produced a specific enhancement of production and/or release of granulocytes in a mammal.

In addition to the naturally occurring allelic forms of hpG-CSF, the present invention also covers other hpG-CSF products, such as polypeptide analogs of hpG-CSF and fragments of hpG-CSF. Following the procedures of the above mentioned published patent application by Alton et al. (WO/83/04053), artificial genes coding for microbial expression of polypeptides having a primary structure different from that herein specified in terms of the identity or location of one or more residues (e.g., substitutions, terminal and intermediate additions and deletions) can be easily designed and produced. Alternately, modifications of cDNA and genomic genes may be readily performed by well known site-directed mutagenesis techniques and used to generate their analogs and derivatives. Such products would share at least one of the biological properties of hpG-CSF but may differ in others. For example, projected products of the present invention include those which are shortened by deletion, etc., those which are more stable to hydrolysis (and, therefore, may have more pronounced or longer lasting effects than the naturally occurring factor), those which have been altered to delete one or more of potential sites for o-glycosylation (which may result in higher activities for yeast-produced products); or those which have one or more cysteine residues deleted or substituted by alanine or serine residues, etc. and potentially can be more easily isolated in the active form from microbial systems; or those which have one or more tyrosine residues substituted by phenylalanine and may bind more or less readily to hpG-CSF receptors on target cells. Polypeptide fragments duplicating only a part of the continuous amino acid sequence or secondary structure within hpG-CSF are also covered. Such fragments may possess one activity (e.g., receptor binding) but not others (e.g., colony-stimulating activity). It is noteworthy that an activity is not necessary for any one or more of the products of the present invention to have a therapeutic utility [see Weiland et al., Blut, 44, 173-175 (1982)] or a utility in other contexts, such as in analysis of hpG-CSF antagonism. Competitive antagonists may be very useful, for example, in cases of overproduction of hpG-CSF.

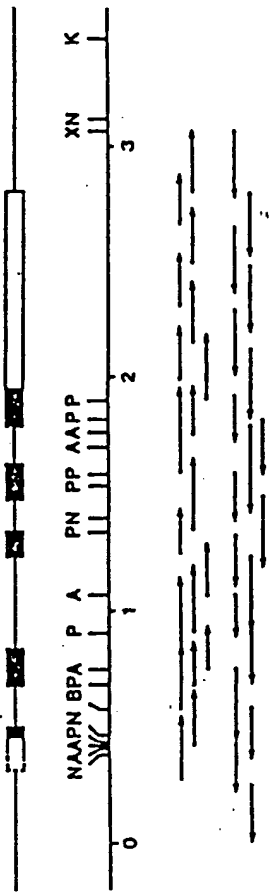
According to another aspect of the present invention, the DNA sequence described herein which encodes the hpG-CSF polypeptide is valuable in providing the information on the amino acid sequence of the mammalian protein which has heretofore been unavailable despite analytical processing of isolates of naturally occurring products. The DNA sequences are also highly valuable, allowing for large-scale microbial production of hpG-CSF by a variety of recombinant techniques. In addition, the DNA sequences provided by the present invention are

useful in generating new and useful viral and circular plasmid DNA vectors, new and useful transformed and transfected microbial prokaryotic and eukaryotic host cells (including bacterial and yeast cells and mammalian cells grown in culture), and new and useful methods for growth, by culture, of such microbial host cells capable of expression of hpG-CSF or its related products. The DNA sequences of the present invention are also highly suitable materials for use as labeled probes in isolating hpG-CSF and related protein-encoding human genomic DNA as well as cDNA and genomic DNA sequences of other mammalian species. The DNA sequences may also be useful for various alternative methods of protein synthesis (e.g., in insect cells) or in genetic therapy in humans and other mammals. The DNA sequences of the present invention are expected to be useful in developing transgenic mammalian species that may serve as eukaryotic "hosts" for the production of hpG-CSF and other hpG-CSF products in large quantities [for general information, see Palmiter et al., Science, 222(4625), 809-814 (1983)].

Application examples of the hpG-CSF fragments and polypeptide analogs of the present invention include the reports on the immunological activities of synthetic peptides that substantially duplicate the amino acid sequence in naturally occurring proteins, glycoproteins and nucleoproteins. More specifically, relatively low molecular weight polypeptides have been shown to participate in immune reactions which are similar in duration and extent to the immune reactions of physiologically significant proteins such as viral antigens, polypeptide hormones, and so on. The immune reactions of such polypeptides include provocation of the formation of specific antibodies in immunologically active animals [Lerner et al., Cell, 23, 309-310 (1981); Ross et al., Nature, 294, 654-656 (1981); Walter et al., Proc. Natl. Acad. Sci. (USA), 77, 5197-5200 (1980); Lerner et al., Proc. Natl. Acad. Sci. (USA), 78, 3403-3407 (1981); Walter et al., Proc. Natl. Acad. Sci. (USA), 78, 4882-4886 (1981); Wong et al., Proc. Natl. Acad. Sci. (USA), 78, 7412-7416 (1981); Green et al., Cell, 28, 477-487 (1982); Nigg et al., Proc. Natl. Acad. Sci. (USA), 79, 5322-5326 (1982); Baron et al., Cell, 28, 395-404 (1982); and Lerner, Scientific American, 248, No. 2, 66-74 (1983)]. Also see Kaiser et al. [Science, 223, 249-255 (1984)] for the biological and immunological activities of synthetic peptides that roughly share the secondary structures of peptide hormones but may not share their primary structures.

Although the present invention has been described in terms of its preferred embodiments, it is understood that variations and modifications can be achieved by those skilled in this field. Therefore, the appended claims are intended to cover all such equivalent variations that come within the scope of the invention as claimed.

FIG.1



国 際 調 査 報 告

International Application No. PCT/US86/01780

IPC[4] C12P 19/00; C12P 19/34; 21/02; C07K 3/00; A61K 37/02

U.S. 435/68, 70, 91, 172.3, 253, 255, and 317; 514/12; 536/27; 530/351; 935/13, 11, 32

CHEMICAL ABSTRACTS DATA BASE (CAS), 1967-1986
BIOLOGICAL ABSTRACT DATA BASE (Biosis) 1969-1986
Key Words: Colony Stimulating Factor; human; gene; vector

Category	Citation of Document	Relevant to Claim No.
Y	BLOOD, (New York) (NICOLA ET AL) issued September 1979 "Separation of functionally distinct human granulocyte-macrophage colony-stimulating factors", see pages 614-627.	1-44
Y	PROCEEDINGS NATIONAL ACADEMY OF SCIENCES (Washington, D.C., USA), (WELTE ET AL), Volume 82, issued March, 1985, "Purification and biochemical characterization of human pluripotent hematopoietic colony-stimulating factor", see pages 1526-1530.	1-44
Y	PROCEEDINGS NATIONAL ACADEMY OF SCIENCES (Washington, D.C., USA), (ZALLER-McFARLAND ET AL), Volume 79, issued November, 1982, "Oligonucleotide-directed mutagenesis as a general and powerful method for studies of protein function", see pages 6409-6413.	38-43

Date of the Actual Completion of the International Search: 10 November 1986

Date of Mailing of this International Search Report: 19 NOV 1986

International Searching Authority: ISA/US

Signature of Authorized Officer: Stephanie Seidman Ph.D., J.D.

PCT/US86/01780

Category	Citation of Document	Relevant to Claim No.
Y	J. BIOLOGICAL CHEMISTRY, (Bethesda, Md.), (HEWICK ET AL), Volume 256, issued 10 August 1981, "A Gas-Liquid Solid Phase Peptide and Protein Sequenator", see pages 7990-7997.	1-34,44
Y	NATURE, (London, England), (GOUGH ET AL), Volume 309, issued June 1984, "Molecular cloning of cDNA encoding a murine haematopoietic growth regulator, granulocyte-macrophage colony stimulating factor", see pages 763-767.	1-34,44
Y	PROCEEDINGS NATIONAL ACADEMY OF SCIENCES, (Washington, D.C. USA), (NICOLA ET AL), Volume 81, issued June 1984, "Binding of the differentiation-inducer, granulocyte-colony-stimulating factor, to responsive but not unresponsive leukemic cell lines", see pages 3765-3769.	35-37
Y	PROCEEDINGS NATIONAL ACADEMY OF SCIENCES, (Washington, D.C., USA), (LERNER ET AL), Volume 78, issued June, 1981, "Chemically synthesized peptides predicted from the nucleotide sequence of the hepatitis B virus genome elicit antibodies reactive with the native envelope protein of Dane particles", see pages 3403-3407.	33-33
Y	MOLECULAR AND CELLULAR BIOLOGY, (Washington, D.C. USA), (OKAYAMA ET AL), Volume 3, issued February, 1983, "A cDNA Cloning Vector that Permits Expression of cDNA Inserts in Mammalian Cells", see pages 280-289.	1-44

continuation from 1st page **Intra-bureau Registration No.**

C 07 K 13/00	8318-4H
C 12 N 1/16	K-6712-4B
1/20	C-6760-4B
5/00	B-6760-4B
15/00	A-8412-4B
//(C 12 P 21/02	
C 12 R 1:19)	
(C 12 P 21/02	
C 12 R 1:865)	
C 12 P 21/02	
C 12 R 1:91)	